# Water Transmission Line Leak Detection using Extended Kalman Filtering

A Thesis Submitted to

the College of Graduate Studies and Research

in Partial Fulfillment of the Requirements

for the Degree of Master of Science

in the Department of Mechanical Engineering

University of Saskatchewan

Saskatoon

By

Ryan M. Lesyshen

Spring 2005

# Permission to Use

In presenting this thesis in partial fulfillment of the requirements for a postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that the permission for copying this thesis in any manner, in whole or in part for scholarly purposes, may be granted by the professors who supervised my thesis work or, in their absence, by the Head of the Department or Dean of the College in which my thesis work was conducted. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis. Requests for permission to copy or to make other use of material in this thesis, in whole or part, should be addressed to:

Head of the Department Mechanical Engineering

University of Saskatchewan

Engineering Building

57 Campus Drive

Saskatoon, Saskatchewan S7N 5A9

Canada

# Abstract

A model-based estimation process is implemented in simulation of a water transmission line for the purpose of leak detection. The objective of this thesis is aimed at determining, through simulation results, the effectiveness of the Extended Kalman Filter for leak detection.

Water distribution systems often contain large amounts of unknown losses. The range in magnitude of losses varies from 10 to over 50 percent of the total volume of water pumped. The result is a loss of product, including water and the chemicals used to treat it, environmental damage, demand shortfalls, increased energy usage and unneeded pump capacity expansions. It is clear that more control efforts need to be implemented on these systems to reduce losses and increase energy efficiencies. The problems of demand shortfalls, resulting from lost product, are worsened by the limited availability of water resources and a growing population and economy. The repair of leakage zones as they occur is not a simple problem since the vast majority of leaks, not considered to be major faults, go undetected.

The leak detection process described in the work of this thesis is model based. A transient model of a transmission line is developed using the Method of Characteristics. This method provides the most accurate results of all finite-difference solutions to the two partial differential equations of continuity and momentum that describe pipe flow. Simulations are run with leakage within the system and small transients are added as random perturbations in the upstream reservoir head. The head measurements at the two pipe extremes are used as inputs into the filter estimation process.

The Extended Kalman Filter is used for state estimation of leakage within the transmission line. The filter model places two artificial leakage states within the system. The estimates of these "fictitious" leakage states are then used to locate the actual position and magnitude of leakage within the transmission line. This method is capable of locating one leak within the line.

The results of the Extended Kalman Filter (EKF) process show that it is capable of locating the position and magnitude of small leaks within the line. It was concluded that the EKF could be used for leak detection, but field tests need to be done to better quantify the ability of these methods. It is recommended that a multiple filtering method be implemented that may be able to locate the occurrence of multiple leakage.

# Acknowledgements

The author expresses his gratitude to his supervisors, Dr. Saied Habibi and Dr. Richard Burton, for their guidance and support during the course of this research and the writing of this thesis. The technical help, guidance and hospitality of Dr. Bryan Karney and his University of Toronto research team are also gratefully acknowledged.

The patience, support and understanding of the author's wife, Katherine Lesyshen, have been invaluable during the course of this research.

# Table of Contents

# List of Figures

**Chapter 2**



**Chapter 3**



**Chapter 4**



**Chapter 5**



**Appendix A**



**Appendix B**



# List of Tables

# Nomenclature

## Pipe Modeling Nomenclature

| | |
|---|---|
| $A$ | Pipe cross-sectional area [$m^2$] |
| $a$ | Wave speed [$m/s$] |
| $B$ | Pipe wave velocity constant |
| $C^+, C^-$ | Positive and negative characteristic equation sets |
| $c_1$ | Pipe loading condition |
| $D$ | Pipe diameter (inner) [$m$] |
| $D/Dt$ | Total derivative |
| $E$ | Young's modulus [$N/m^2$] |
| $e$ | Wall roughness of pipe [$mm$] |
| $f$ | Darcy-Weisbach friction factor |
| $g_x$ | Acceleration due to gravity [$m/s^2$] |
| $H$ | Hydraulic head [$m$] |
| $P$ | Pressure [$Pa$] |
| $Q$ | Flow rate [$m^3/s$] |
| $R$ | Pipe friction constant |
| $R_1$ | Pipe Radius [$m$] |
| $r$ | Radial pipe position [$m$] |
| $u_r$, $u_\theta$, $u_x$ | Radial, rotational and axial components of fluid velocity [$m/s$] |
| $u$ | Fluid velocity (axial) [$m/s$] |
| $V$ | Fluid velocity [$m/s$] |

| | |
|---|---|
| $x$ | Distance along length of pipe [$m$] |
| $z$ | Pipe elevation [$m$] |
| $\beta$ | Bulk modulus of water [$N/m^2$] |
| $\varepsilon_T$ | Circumferential strain |
| $\lambda$ | Unknown multiplier used in Method of Characteristics derivation [$m/s$] |
| $\mu$ | Fluid viscosity [$Ns/m^2$] |
| $\rho$ | Fluid density [$kg/m^3$] |
| $\sigma_\theta$ | Circumferential pipe stress [$N/m^2$] |
| $\sigma_z$ | Axial pipe stress [$N/m^2$] |
| $\tau$ | Shear Stress [$N/m^2$] |
| $\upsilon$ | Poisson's ratio |

**State and Estimation Nomenclature**

| | |
|---|---|
| $\mathrm{cov}[X,Y]$ | Covariance of the random processes $X$ and $Y$ |
| $E[X]$ | Expectation of a random process $X$ |
| $e_k^-$ | *A priori* error estimate |
| $e_k$ | *A posteriori* error estimate |
| $f[x_k]$ | Non-linear state function |
| $G_k$ | Input matrix |
| $H_k$ | Output matrix that linearly connects outputs and states |
| $h[x_k]$ | Non-linear output matrix |

| | |
|---|---|
| $I$ | Identity matrix |
| $J_x$ | Jacobian of $f[x_k]$ evaluated at $\hat{x}_k$ |
| $J_h$ | Jacobian of $h[x_k]$ evaluated at $\hat{x}_{k+1}^-$ |
| $K_k$ | Kalman Gain at time $t_k$ |
| $P_k^-$ | Unrefined (*a priori*) estimate of covariance matrix at time $t_k$ |
| $P_k$ | Refined (*a posteriori*) estimate of covariance matrix at time $t_k$ |
| $Q_k$ | System noise covariance matrix |
| $R_k$ | Measurement noise covariance matrix |
| $u_k$ | Input vector |
| $v_k$ | Vector of white measurement noise |
| $w_k$ | Vector of white system noise |
| $x_k$ | State vector at time $t_k$ |
| $\hat{x}_k^-$ | Predicted estimate of the state vector at time $t_k$ |
| $\hat{x}_k$ | Refined estimate of the state vector at time $t_k$ |
| $z_k$ | Vector of defined measurements |
| $\phi_k$ | System or transition matrix of constants for time $t_k$ |

# Chapter 1: Introduction

## 1.1 Preliminary Remarks

Water is our most precious natural resource. It is vital to the environment, the economy and to human health. The distribution systems that bring water from its source to the tap are therefore just as vital to economic and human survival. Maintaining good health within the distribution system is an important and difficult job. Loss monitoring within water distribution systems is the focus of this thesis.

Water distribution systems are buried infrastructure, and therefore it is extremely difficult to locate problem areas. Pipeline leakage has a number of negative effects, including the loss of water and the chemicals used to treat it, an increased use of energy, possible drinking water contamination due to pipe infiltration, demand shortages leading to unnecessary capacity expansions and the possibility for environmental damage [1].

Pipeline leakage represents a significant portion of the water leaving a supply pumping station. In Europe unaccounted for water is typically in the order of 9-30% of the total volume pumped [2]. A recent study by Brothers states that in North America, some utilities record water losses in the order of 20-50% [3]. Generally, losses increase with the size and age of water distribution systems. According to Bullee, new municipal water systems are known to lose ten to fifteen percent of their total production [4]. This number increases for older systems; many large cities report 40-50% losses within their distribution systems. The municipal water systems of Malaysia and Bangladesh record 43% and 56% losses respectively [2] [5]. In many situations where demand shortages occur, supply shortfalls may easily be made up by reducing losses due to leakage.

Large leakage or blow outs are easy to locate because they often show themselves through surfaced water and system pressure loss resulting in inadequate supply past the location of the leak. Smaller leaks that do not create surface ponds or large system pressure losses are much harder to detect and may only be found through adequate system monitoring.

It is evident that better system monitoring is required to detect leaks at early onset. As will be presented in the review of the literature, several attempts have been made to develop methodologies or devices to do this with some limited success. However, it is clear that the problem is a complex one and poses many challenges. This challenge, then, was one of the motivations for this study.

## 1.2 Project Origin

### 1.2.1 The State of the World's Water

In order to understand the exigency of leak detection methods within water distribution systems, it is important to first define the state of the world's fresh water supply. Fresh water lies at the fundamental roots of food production, and therefore human existence. Historically, the fall of many great empires was directly related to either a shortage of fresh water or an increase in ground and water salinity resulting in the collapse of crop production [6].

It is estimated that about 2.78% of the world's water is fresh. Out of this small fraction, 99.357% of it is trapped inside glaciers. The other 97% of water available is saline and cannot be consumed in its present state [7]. Fresh water is also geographically unbalanced, and therefore many parts of the world are said to be "water strained". A few of the more seriously strained areas include India, Africa, the Middle East, China, and many areas of the United States. The number of geographical regions referred to as "water strained" are increasing as both industry and human consumptive needs increase.

Poor or inefficient use of available water sources inevitably leads to a search for new sources. In the past, large public works projects built dams to meet the supply needs of the growing population and industry. According to Postel, the rapid dam construction of the past century is largely responsible for the tripling of global water use since 1950 [6]. Today's society, consisting of an expanding population and industry, would not have been possible without the past centuries' obsession with dam construction. But dam construction has dropped drastically due to public pressure over environmental issues, large capital costs and the fact that there are very few new places available to build dams.

In today's society, where the capacity of dams has neared the maximum value, 26 percent of fresh water used in the United States comes from underground reservoirs, which equates to 315.3 million cubic meters per day [8]. This rate of ground water removal by far exceeds the rate of replenishment. In the American Midwest, the Ogallala aquifer, which is North America's largest body of water, is being mined 15 times faster than its natural replenishment rate [9]. It is estimated that 71.9 million cubic meters of water per day is extracted from this reservoir; that is 23 percent of all US mined water and just under 5 percent of the total water use within the United States. Some have estimated that more than half of the Ogallala aquifer is already gone, and others predict that within 10 to 20 years "humans of the High Plains will be staring down tens of thousands of dry holes" [6] [10]. The situation occurring over the Ogallala aquifer is by no means an anomaly. The vast majority of the worlds fresh water aquifers are in a non-sustainable state [6].

Across the globe, continents are drying as the ground covering exploited aquifers subsides, resulting in a permanent loss of aquifer capacity. Mined water inevitably reaches the sea through the hydrological cycle. An imbalance therefore exists between continental inflow and outflow as deep ground water is being surfaced and inevitably drained into the sea. It has been estimated that "the continents are losing 1800 billion cubic meters of fresh water a year. If this trend continues, over the next 100 years, the continents will lose 180,000 billion cubic meters of fresh water, which is approximately equivalent to the total volume of water within the hydrological cycle" [9].

Water shortages have resulted in the growth of new technologies, as heavy consumers such as the farmers of the Ogallala area and the people of the other water strained regions become forced conservationists. Some of these technologies include drip irrigation, desalination and leak detection methods. Other effective conservation methods have included increased governmental management, public awareness, water marketing and increased water pricing. In areas where water scarcity occurs, such as the Middle East, most if not all of these methods are utilized depending on their geo-political merit.

### 1.2.2 Aging Water Distribution Systems

Water distribution infrastructure is generally in poor condition and is deteriorating rapidly [11]. Due to the poor state of infrastructure and the lack of capital funds, many municipal engineers manage crisis situations instead of healthy preventative maintenance schedules. Rebuilding aging water distribution systems will be this generation of municipal engineer's largest challenge. To illustrate this point, several specific examples are considered. The data cited below represent average values taken from three municipalities in the province of Quebec, namely Chicoutimi, Gatineau, and Saint-Georges. These municipalities were used in an infrastructure study by Pelletier *et al.* due to the availability of computerized records [11].

Water distribution infrastructure in North America may be broken into four age categories or "bubbles" that correspond to four urbanization periods [12]. Modern central water distribution was introduced around 1850. It is estimated that approximately 5 percent of current water distribution infrastructure was built between 1850 and 1940. The next large "bubble" of infrastructure occurred during the rapid urban growth period following the Second World War. Approximately 25 percent of current infrastructure was built between 1940 and 1960. The urbanization period between 1961 and 1975 added the largest percentage of current infrastructure, at approximately 40 percent. According to Pelletier *et al.*, due to the use of poorer quality materials and poor installation techniques, pipes of this period exhibit the highest rate of failure [11]. The remaining 30 percent of current infrastructure may be considered new and exhibits the lowest rate of failure. The age group of this infrastructure is from 1975 to present day.

The type of materials used within distribution systems are closely related to the urbanization period in which they were laid [11]. Cast iron and asbestos-cement pipes were used within the first two "bubbles" of infrastructure with cast iron representing the majority. Cast and ductile iron were used in the third "bubble". The newest infrastructure or forth "bubble" represents the period in which PVC pipes were introduced. PVC pipes have the lowest breakage rate at approximately 0.02 [breaks/km/year], ductile iron the middle at approximately 0.2 [breaks/km/year] and cast

iron exhibits the greatest number of pipe breaks at approximately 0.6 [breaks/km/year] [11]. PVC is pressure fitted at the joints and does not exhibit the leaky joint properties of the older methods.

According to a survey conducted by the National Research Council of Canada, each year an average of 35 pipe failures occurs per every 100 kilometers of pipeline [13]. The three municipalities, cited above, report a similar average failure rate of 34 pipe breaks per 100 kilometers of pipeline per year. The three oldest "bubbles" of infrastructure, within these municipalities, have failure rates of approximately 54, 48 and 52 pipe breaks per 100 kilometers of pipeline per year, respectively. The newest "bubble" of infrastructure, 1975 to the present, has a breakage rate of approximately 10 pipe breaks per 100 kilometers of pipe per year.

The statistics stated above point out a glaring problem. Seventy percent of pipes within water distribution systems, installed between 1850 and 1975, display a high breakage rate of about 50 pipe breaks per kilometer each year. This massive amount of aging, leaky infrastructure represents a huge challenge that is quickly "getting out of hand" [11]. In order to effectively schedule pipe replacement and maintenance, adequate information is needed. Real time leak monitoring would serve as a vital tool for municipal managers and engineers.

### 1.2.3 Transient Pipeline Analysis and Leak Detection

One approach (of particular interest in this research) that has been used to detect leakage involves gaining unknown system information through dynamic system monitoring. Information from the system may be obtained by "listening" to fluid transients. Transients are unsteady fluid flow, where fluctuations in system pressures and flows are brought about by various external forces. Some causes of transient flow include valve operations, pump startup and shutdown, reservoir waves and pipe failures. Transient waves are shaped by their surroundings. The speed of the transient wave depends on pipe characteristics while the shape of the wave is influenced by the various devices that exist within the pipeline. Therefore, each unique water distribution system will have a unique

transient behaviour that depends upon the various devices within the system. Some of these devices include valves, orifices, elbows, branches and pipe breaks. If a system could be effectively modeled with transient equations, control logic could be used to compare system measurements with modeled outputs in order to determine the possible location of leakage within the system.

Transients may be modeled by various techniques. All methods of solution start with the general equations of momentum, continuity and energy, but their solution differs in how the partial differential equations are handled due to differences in the simplifying assumptions [14]. Some methods are not conducive to large numerical computer implementation; these methods include: graphical, arithmetic, implicit and linear analyzing methods. Three methods that are more readily adapted into large computer models include: method of characteristics, finite-difference method, and the finite element method. Of these three methods, the method of characteristics has been cited by many as superior [14] [15] [16] [17]. Some of its advantages include ease of programming, ability to handle complex systems, stability criteria are firmly established, correctly simulates steep wave fronts and has the best accuracy of any finite-difference method [14] [15].

There are a number of different technologies that have been applied towards leak detection in water distribution systems. Leak detection techniques that have been proposed in the literature include: ground penetrating radar or infrared spectroscopy [18]; acoustic methods [19]; reflected wave or timing methods [20]; transient damping methods [21]; frequency response methods [22]; pressure and flow deviation methods [23] [24]; mass and volume balance methods [23] [25]; generalized likelihood methods [26]; optimal valve regulation methods [27]; genetic algorithm methods [28] [29]; and system identification and state estimation techniques [30] [31] [32] [33] [34] [35] [36]. The first two methods that are mentioned do not use system models while the rest are model-based techniques. A few of the model-based methods are expanded upon below.

Brunone, in 1999, used the timing of reflected pressure waves to determine the location of leakage within outfall pipes [20]. He compared pressure traces, taken at the upstream end of the pipe, from new outfall pipes with those which had suspected leakage. The occurrence of transient damping determined the presence of a leak and the timing of the damping determined its location. Wang *et al.*, in 2002, studied the damping of fluid transients in order to determine the magnitude and location of leakage within a pipeline [21]. They noted that damping due to pipe friction was equal for all harmonic components of the transient, but leak damping effects were different for each harmonic component. This fact was used to detect the location and magnitude of leakage within a simple system. Due to the complex wave forms that are created by various network components, this method cannot be used for pipe networks.

Mpesha *et al.*, in 2001, used the pipelines' frequency response to determine the position of leakage within a few series and parallel open loop systems [22]. The transfer matrix method was used to simulate steady-oscillatory flow. Simulation results, with and without leakage, were compared. In a real system, steady oscillatory flow would be set up through controlled valve operations and pressure and flow fluctuations would be recorded. This procedure would be repeated for a number of valve fluctuating frequencies. The frequency response of the real system could then be determined and compared with an expected computer simulation in an attempt to determine discrepancies due to pipe leakage.

Liou and Tian, in 1995, developed a model for a single pipeline using transient flow simulations and compared its results with field data [24]. They used the Cauchy and time-marching algorithms to determine discrepancy patterns between the modeled and the measured real-time values of pressure and flow at the ends of the pipe. They noted that the presence of noise made it difficult in many situations to detect leakage. Liou, in 1994, proposed a leak detection method based on the mass balance principle that "what goes in must come out" [25]. A steady state model was used in this work.

Pudar and Liggett, in 1992, solved the inverse problem using measurements of pressure and/or flow [37]. The problem was formulated by modeling unknown leakage with the orifice equation at discrete locations. Leakage was estimated using a least squares regression. A better prediction of leaks was made by Liggett and Li-Chung in 1994 by incorporating large amounts of data for a better determination of the friction factor [38]. The search space, for determining pipeline leakage, may be very large and the solution is dependent on the starting point. Therefore their methods could not guarantee convergence to the global optimum.

Vitkovsky *et al.*, in 2000, used a genetic algorithm in order to solve the inverse problem [28]. Genetic algorithms also do not guarantee a global optimum, but due to the addition of random events within the solution, a larger search space may be covered. A disadvantage of this method is the slow rate of convergence within complex systems. Tang *et al.*, also used genetic algorithms in conjunction with a system model to calibrate and solve for unknown pipe leakage [29]. Again, convergence rates were slow.

System identification and state estimation techniques have been applied to an array of fluidized systems for the purpose of fault detection. State estimation techniques, in conjunction with a process model, may be used to monitor or determine immeasurable variables such as system states and parameters. Both Willsky and Isermann present a variety of estimation techniques [39] [40]. The most frequently used method, which is of special interest within this thesis, is the Kalman or Extended Kalman Filter. Two advantages of these methods are fast convergence and a robust ability at handling noise.

Fault detection, through Kalman Filtering and other estimation techniques, is most widely used within process operations in chemical plants. This is due to their need for quick, online diagnosis of process faults that lead to unsafe situations or environmental damage. Dalle Molle and Himmelblau, in 1987, used the Kalman Filter for fault detection within a single stage evaporator [35]. Li and Olson, in 1991, applied the Extended Kalman Filter to a closed-loop non-linear distillation process [36]. Chmielewski and Manousiouthakis, in 2000, used the Kalman Filter for loss accounting within a refinery setting [34]. Sun *et*

*al.*., in 2002, applied a least squares estimation approach with a forgetting factor to boiler leak detection in the chemical and refinery industry [31].

The most thorough water distribution leak detection study was published in 1991. Carpentier and Cohen wrote a paper called "State estimation and leak detection in water distribution networks" [30]. In this work, they determined, using graph theory, which variables were observable from available measurements. They looked at improving observability through well placed measurements and solved the least squares problem for leak detection. A real distribution system located in the city of Paris, France, was used in the study. Fifteen flow meters were placed throughout the system. Head measurements were not considered due to low system pressures. Leak detection was run at night since consumption levels were low and the overall effect of the leakage was more noticeable at this time.

Carpentier and Cohen made the following conclusions. An accurate network model was difficult to obtain due to errors in some technical data and little knowledge on the state of valves and other devices within the system. Determining the state of the system led to improved operational knowledge, and improved system state and efficiency. The calibration methods used were not satisfactory but the approximate methods produced satisfactory leak detection results. The choice of instrumentation was also difficult due to the need for high accuracy at an acceptable cost. Leak detection was successfully performed but it was noted that a better method would incorporate "estimating consumption rates every five minutes and analyzing the results using statistical filters to produce a more accurate diagnosis" [30].

Benkherouf and Allidina, in 1988, performed leak detection on a gas pipeline using the Extended Kalman Filter [32]. The proposed method included artificial leak states within the system model which were set as unknowns to be estimated by the filter. The system model used the method of characteristics for the solution of the momentum and continuity equations. Results showed that a very coarse discretization, requiring only modest computational effort, was successful in detecting and locating leakage within the

pipeline. It is noted that this method may only be used to detect a maximum of two leak locations due to the observability of the system [41]. In order to detect multiple leakage both Verde and Digernes proposed the use of a bank of Kalman Filters [41] [42]. Verde was able to detect multiple leakage in a simple fluid line with the use of two head measurements, located at the pipe boundaries.

## 1.3 Research Scope and Objectives

Water distribution systems are characterized by a massive array of pipes and boundary devices. There have been many approaches forwarded for the detection of leaks in pipelines. One approach is that associated with the use of transient information and pipe line models to detect the presence and identify the location of leaks. Due to the many components of pipelines, transients within a network may be reflected by numerous sources, making the problem extremely difficult to characterize. There is also a large degree of noise associated with measurements. The solution method must prove to be robust to noisy systems, thus the filtering techniques used must be robust to noise. The Extended Kalman Filter (EKF) has been used in a wide variety of engineering fields due to its excellent ability of finding stable solutions within the presence of noise. The EKF is tolerant to the presence of extraneous noise, model simplifications and inaccuracies [43]. The quick convergence rate of the EKF method also makes it desirable for leak detection.

The objective of this thesis is to apply, through simulation, the EKF to the problem of leak detection and location within a water transmission line. A model-based leak detection technique is developed and evaluated. The technique requires adequate monitoring of pressure and/or flow as inputs into a state estimation scheme. State estimation through Extended Kalman Filtering is applied to a system model and the performance of the technique is discussed.

Due to the highly complex nature of most water distribution systems, the EKF will be applied to a simple transmission line simulation. The work of this thesis may be seen as a

first step in studying the EKF accuracy in detecting and locating faults. If the EKF performance does not meet expectations for a simple system, then its application towards larger sets of distribution lines may not be advised.

The level of success of these methods is more closely related to the accuracy of the position estimates then to the accuracy of the magnitude estimate. This is due to the fact that excavation is needed in order to fix broken sections. A system capable of detecting leak locations accurate to a few meters would be seen as highly successful since this would limit the amount of digging to a small section. However, estimates with an accuracy of tens of meters may also be perceived as successful since they would greatly reduce the search space on a long transmission line.

The problem will be formulated in such a way that two defined fictitious leakage points will be placed along the length of the transmission line model. These fictitious leakages will be estimated by the EKF. An estimate of the actual leak magnitude is given by the continuity equation that states that it must equal the sum of the two fictitious leaks. The estimated location is given by a linear interpolation of the two fictitious leaks.

## 1.4 Thesis Outline

Chapter 2 presents the theory of transient fluid modeling and the method of characteristics, which is the solution method for the partial differential equations of motion and continuity. Chapter 3 presents the theory of the Kalman and Extended Kalman Filter. Chapter 4 applies the transient theory of Chapter 2 and the Kalman Filter of Chapter 3 to a transmission line model. In Chapter 5 the EKF is applied to the transmission model simulation and the results are presented. Conclusions and recommendations for future work are given in Chapter 6.

# Chapter 2: Transient Pipe Flow Equations

## 2.1 Preliminary Remarks

Transient pipe flow is fully described by the momentum and continuity equations. There are two dependent variables in the momentum and continuity equations - flow and pressure - and there are two independent variables - time and space. The objective is to determine a value for the dependent variables at a prescribed time and position in space. Since transients vary both spatially and temporally, they must be described by partial differential equations. The following chapter contains a derivation of the momentum and continuity equations.

The final forms of the derived equations are analyzed. Simplifying assumptions are made and explained in order to form generalized equations that can be more easily solved. These generalized equations are given by the well known method of characteristics [14]. The method of characteristics converts the two partial differential equations of momentum and continuity into four ordinary differential equations. These four equations are presented in a discretized form at the end of the chapter.

## 2.2 The Equations of Transient Pipe Flow

### 2.2.1 Transient Momentum Equation

In this section the momentum equation is derived from the Navier-Stokes equations. Figure 2.1 displays the pipe geometry and defines the parameters of the problem.

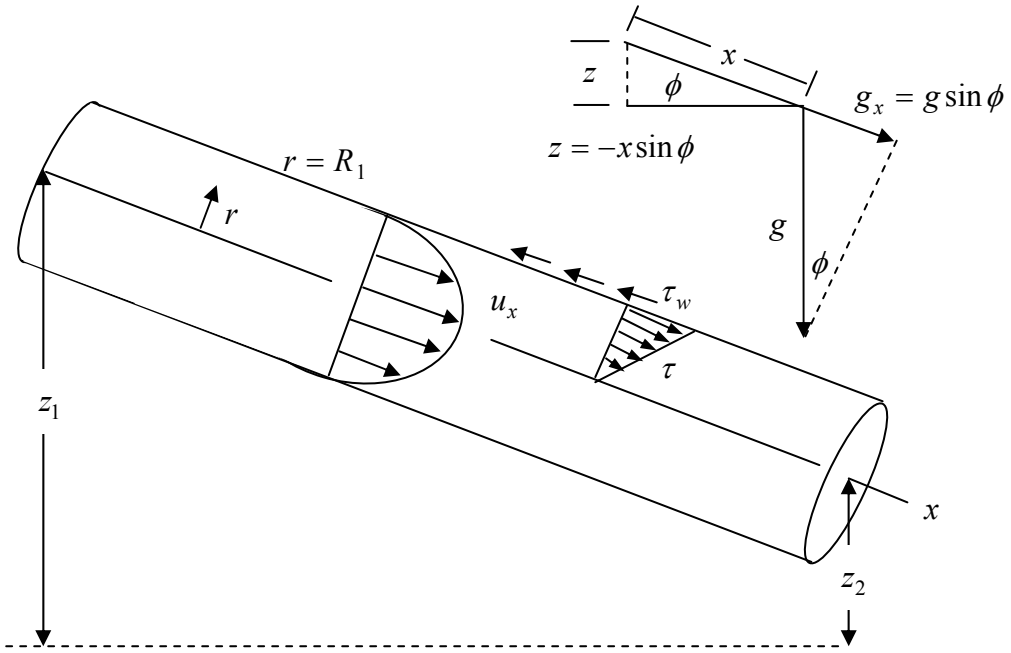


**Figure 2. 1: Pipe Geometry**

All fluid flows are governed by the Navier-Stokes equations. Pipeline fluid flow is essentially one-dimensional and therefore only the *x*-direction equation of the Navier-Stokes equation is needed. This equation in cylindrical coordinates is given as:

$$\rho\left(\frac{\partial u_x}{\partial t}+u_r\frac{\partial u_x}{\partial r}+\frac{u_\theta}{r}\frac{\partial u_x}{\partial \theta}+u_x\frac{\partial u_x}{\partial x}\right)=\rho g_x-\frac{\partial P}{\partial x}+\mu\left(\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u_x}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2 u_x}{\partial \theta^2}+\frac{\partial^2 u_x}{\partial x^2}\right),[2.1]$$

where the velocities components are given as radial, $u_r$, rotational, $u_\theta$, and axial, $u_x$, $\rho$ is fluid density, $g_x$ is axial body force per unit mass, $P$ is pressure and $\mu$ is fluid viscosity. Assuming that the fluid is Newtonian, one-dimensional, incompressible and has constant properties:

$$u_r = 0,$$
$$u_\theta = 0,$$
$$u_x = u(r).$$

For ease, the velocity will simply be denoted as $u$. The Navier-Stokes equation now becomes:

$$\rho\left(\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}\right)=\rho g_x-\frac{\partial P}{\partial x}+\frac{\mu}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right), \qquad [2.2]$$

where the terms on the left hand side represents the inertial or transient effects of the flow and convective acceleration. On the right hand side, the terms represent gravity forces, pressure forces and shear forces respectively. Noting that shear stress is given as:

$$\tau=\mu\frac{\partial u}{\partial r}, \qquad [2.3]$$

equation [2.2] can be re-written as:

$$\rho\left(\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}\right)=\rho g_x-\frac{\partial P}{\partial x}+\frac{1}{r}\frac{\partial(r\tau)}{\partial r}. \qquad [2.4]$$

In general shear stress, $\tau$, is a function of viscosity, $\mu$, density, $\rho$, wall roughness, *e*, radial position, *r*, and fluid velocity, *u*. However, viscosity and wall roughness can be assumed constant which leaves $\tau=f(r,u)$. Furthermore, if the effects of frequency-dependent friction are neglected, the shear stress for a given velocity is the same as in steady flow at that velocity [15]. This common practice in transient flow modeling allows two things: evaluation of the shear stress in equation [2.4] as an ordinary differential equation, and modeling of the friction forces using the Darcy-Weisbach equation. Rearranging [2.4] and integrating from the center of the pipe, *r* = 0, to the wall, $r=R_1$ gives:

$$\int_0^{R_1}\frac{d(r\tau)}{dr}dr=\int_0^{R_1}r\left(\rho\frac{\partial u}{\partial t}+\rho u\frac{\partial u}{\partial x}-\rho g_x+\frac{\partial P}{\partial x}\right)dr,$$

$$R_1\tau = \frac{R_1{}^2}{2}\left(\rho\frac{\partial u}{\partial t} + \rho u\frac{\partial u}{\partial x} - \rho g_x + \frac{\partial P}{\partial x}\right),$$

$$\tau = \frac{R_1}{2}\left(\rho\frac{\partial u}{\partial t} + \rho u\frac{\partial u}{\partial x} - \rho g_x + \frac{\partial P}{\partial x}\right). \qquad [2.5]$$

The Darcy-Weisbach equation which relates the shear stress at the wall, $\tau_w$, to a friction factor, $f$, is given as:

$$\tau_w = \frac{\rho f u|u|}{8}. \qquad [2.6]$$

Combining equations [2.5] and [2.6] gives:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} - g_x + \frac{1}{\rho}\frac{\partial P}{\partial x} + \frac{fu|u|}{4R_1} = 0. \qquad [2.7]$$

Body forces, $g_x$, may be represented in terms of elevation, *z*, as shown on Figure 2.1. Since $g_x = g\sin\phi$, equation [2.7] becomes:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} - g\sin\phi + \frac{1}{\rho}\frac{\partial P}{\partial x} + \frac{fu|u|}{2D} = 0, \qquad [2.8]$$

where *D* is the pipe diameter. Equation [2.8] can then be rearranged as:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + \frac{1}{\rho}\frac{\partial}{\partial x}\left(P - \rho g x\sin\phi\right) + \frac{fu|u|}{2D} = 0. \qquad [2.9]$$

Since $x\sin\phi = -z$ equation [2.9] becomes:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + \frac{1}{\rho}\frac{\partial}{\partial x}(P + \rho g z) + \frac{f u|u|}{2D} = 0 \,. \qquad [2.10]$$

Changing fluid velocity to flow rate by substituting $u = Q/A$ gives:

$$\frac{\partial Q}{\partial t} + \frac{Q}{A}\frac{\partial Q}{\partial x} + \frac{A}{\rho}\frac{\partial}{\partial x}(P + \rho g z) + \frac{f Q|Q|}{2DA} = 0 \,, \qquad [2.11]$$

which is the final form of the momentum equation representing transient pipe flow.

### 2.2.2 Continuity Equation

The continuity equation is derived from the principle of mass conservation which states that mass flow in is equal to mass flow out of a control volume. The control volume is represented in Figure 2.2. It consists of three control surfaces which are sections 1 (indicated by $x_1$), 2 (indicated by $x_2$) and the inner walls of the pipe bounded by sections 1 and 2 shown by the dotted lines.

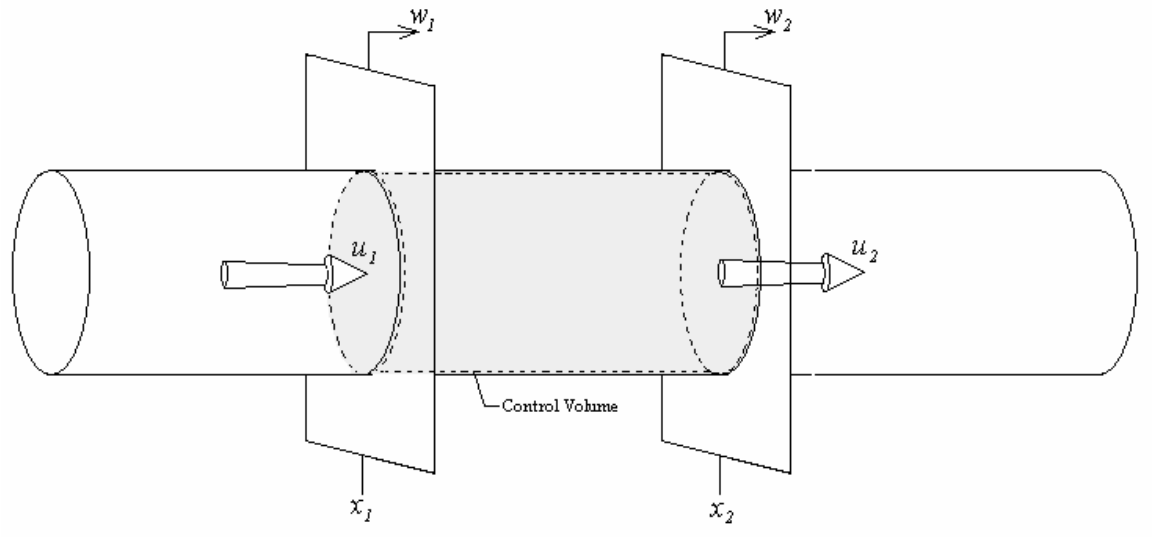


**Figure 2. 2: Control Volume**

Velocity components $u_1$ and $u_2$ represent incoming and outgoing fluid velocities for the control volume. Pressure fluctuations cause expansion and contraction of the pipe walls in the axial direction. Velocities $w_1$ and $w_2$ of control surfaces 1 and 2 compensate for this lateral expansion and contraction. Radial velocities due to expansion and contraction are small and can be neglected. The flow is considered one-dimensional and the sign of a parameter is considered positive if it is in the direction of the flow or in the downstream direction.

Applying the Reynolds transport theorem for the conservation of mass gives:

$$\frac{d}{dt}\int_{x1}^{x2}\rho A dx + \left(\rho A V_s\right)_{out} - \left(\rho A V_s\right)_{in} = 0\,, \qquad [2.12]$$

where $V_s$ is the fluid velocity with respect to the control surface it is crossing. The inlet and outlet control surfaces are subject to expansion and contraction and therefore $V_s$ is a relative fluid velocity given as $V_s = (u_i - w_i)$, where the subscript *i* simply represents the control surface through which the fluid is moving. Therefore equation [2.12] becomes:

$$\frac{d}{dt}\int_{x1}^{x2}\rho A dx + \rho_2 A_2\left(u_2 - w_2\right) - \rho_1 A_1\left(u_1 - w_1\right) = 0\,. \qquad [2.13]$$

The Leibnitz rule can be used to further simplify equation [2.13]. The rule states that:

$$\frac{d}{dt}\int_{f_1(t)}^{f_2(t)} F(x,t)dx = \int_{f_1(t)}^{f_2(t)} \frac{\partial}{\partial t}F(x,t)dx + F\left(f_2(t),t\right)\frac{df_2}{dt} - F\left(f_1(t),t\right)\frac{df_1}{dt}\,,$$

where $f_1$ and $f_2$ are differential functions of *t* and *F(x,t)* and $\partial f/\partial t$ are continuous in space and time [15]. Applying the Leibnitz rule to equation [2.13] gives:

$$\int_{x1}^{x2} \frac{\partial}{\partial t}(\rho A)dx + \rho_2 A_2 \frac{dx_2}{dt} - \rho_1 A_1 \frac{dx_1}{dt} + \rho_2 A_2 (u_2 - w_2) - \rho_1 A_1 (u_1 - w_1) = 0 . \quad [2.14]$$

Since $x_1$ and $x_2$ are fixed to sections 1 and 2, their derivatives with respect to time are the surface velocities $dx_1/dt = w_1$ and $dx_2/dt = w_2$. Therefore, wall velocities cancel out and [2.14] simplifies to:

$$\int_{x1}^{x2} \frac{\partial}{\partial t}(\rho A)dx + \rho_2 A_2 u_2 - \rho_1 A_1 u_1 = 0 . \quad [2.15]$$

According to the mean value theorem the first term on the left side of the equation is given as:

$$\int_{x1}^{x2} \frac{\partial}{\partial t}(\rho A)dx = \frac{\partial}{\partial t}(\rho A)(x_2 - x_1), \quad [2.16]$$

and equation [2.15] becomes:

$$\frac{\partial}{\partial t}(\rho A)(x_2 - x_1) + (\rho A u)_2 - (\rho A u)_1 = 0 . \quad [2.17]$$

Dividing through by $(x_2 - x_1)$ gives:

$$\frac{\partial}{\partial t}(\rho A) + \frac{(\rho_2 A_2 u_2 - \rho_1 A_1 u_1)}{(x_2 - x_1)} = 0 , \quad [2.18]$$

which can be simplified by recognizing that the second term on the left side of the equation, $\frac{\Delta(\rho A u)}{\Delta x}$, is nothing more than the partial derivative of $(\rho A u)$ with respect to $x$. Therefore equation [2.18] becomes:

$$\frac{\partial}{\partial t}(\rho A)+\frac{\partial}{\partial x}(\rho A u)=0\,. \qquad [2.19]$$

Expanding [2.19] by partial fractions gives:

$$\frac{\partial}{\partial t}(\rho A)+u\frac{\partial}{\partial x}(\rho A)+\rho A\frac{\partial u}{\partial x}=0\,. \qquad [2.20]$$

For ease of computation, the concept of a total derivative is introduced. The total derivative of a function $F$ that varies spatially and temporally, $F = f(x,t)$, is represented as:

$$\frac{D}{Dt}F=\frac{\partial F}{\partial x}\frac{\partial x}{\partial t}+\frac{\partial F}{\partial t}\,, \qquad [2.21]$$

where $\frac{D}{Dt}F$ represents the total derivative of function $F$. Noting that $\partial x/\partial t = u$, equation [2.21] becomes:

$$\frac{D}{Dt}F=u\frac{\partial F}{\partial x}+\frac{\partial F}{\partial t}\,. \qquad [2.22]$$

The first two terms of [2.20] represent the total derivative of $(\rho A)$ with respect to the fluid velocity $u$, therefore:

$$\frac{D}{Dt}(\rho A)+\rho A\frac{\partial u}{\partial x}=0\,. \qquad [2.23]$$

Representing total derivatives of $\rho$ and $A$ by $\dot{\rho}$ and $\dot{A}$, and dividing through by $\rho A$, equation (19) can be expanded by partial fractions as:

$$\frac{1}{\rho A}\left(\dot{\rho} A + \rho \dot{A}\right) + \frac{\partial u}{\partial x} = 0,$$

or:

$$\frac{\dot{\rho}}{\rho} + \frac{\dot{A}}{A} + \frac{\partial u}{\partial x} = 0. \qquad [2.24]$$

**Elastic Theory**

Elastic theory is used to transform equation [2.24] into a useful equation relating pressure to flow. The fluid's bulk modulus of elasticity, $\beta$, may be given as:

$$\beta = \rho_o \left(\frac{\partial P}{\partial \rho}\right)_T, \qquad [2.25]$$

which may also be represented by:

$$\frac{\dot{\rho}}{\rho} = \frac{\dot{P}}{\beta}. \qquad [2.26]$$

The pipe wall expansion per unit area, per unit time $\dot{A}/A$ is related to the circumferential strain, $\varepsilon_T$, by the following expression:

$$\frac{\dot{A}}{A} = 2\dot{\varepsilon}_T. \qquad [2.27]$$

Stress and strain in a pipe are related by Hooke's law as:

$$\dot{\varepsilon}_T = \frac{1}{E}\left(\dot{\sigma}_\theta - \upsilon \dot{\sigma}_z\right), \qquad [2.28]$$

where $E$ denotes Young's modulus of the pipe, $\sigma_\theta$ is the circumferential stress, $\sigma_z$ is the axial stress, and $\upsilon$ is Poisson's ratio. Combining equations [2.26], [2.27], and [2.28] with equation [2.24] yields:

$$\frac{\dot{P}}{\beta}+\frac{2}{E}\left(\dot{\sigma}_\theta-\upsilon\dot{\sigma}_z\right)+\frac{\partial V}{\partial x}=0\,. \qquad [2.29]$$

The circumferential stress, in equation [2.29], is related to pressure by:

$$\sigma_\theta=\frac{Pr}{e},$$

therefore,

$$\dot{\sigma}_\theta=\frac{\dot{P}D}{2e}, \qquad [2.30]$$

where $e$ is the thickness of the pipe walls. Axial stresses, in equation [2.29], are given below for the three different loading conditions. These three conditions, given by Chaudhry, are represented by [15]:

Condition 1. Pipe anchored at upstream end only: $\dot{\sigma}_z=\frac{\dot{P}D}{4e}$,

Condition 2. Pipe anchored against axial movement: $\dot{\sigma}_z=\upsilon\dot{\sigma}_\theta$, [2.31]

Condition 3. Pipe anchored with expansion joints throughout: $\dot{\sigma}_z=0$.

Substituting equations [2.30] and [2.31] into [2.29] yields:

$$\frac{D}{Dt}P+\rho a^2\frac{\partial V}{\partial x}=0, \qquad [2.32]$$

where,

$$a^2 = \frac{\beta/\rho}{1+\left[\left(\beta/E\right)\left(D/e\right)\right]c_1}. \qquad [2.33]$$

$a$ is the wave speed of the fluid within the pipe, which is a function of fluid and pipe properties only. The constant, $c_1$, is different for each loading condition. $c_1$ is given as:

Condition 1: $c_1 = 1 - \upsilon/2$,

Condition 2: $c_1 = 1 - \upsilon^2$,

Condition 3: $c_1 = 1$.

Finally, expanding the total derivative and expressing in terms of flow, the continuity equation may be written as:

$$\frac{Q}{A}\frac{\partial P}{\partial x} + \frac{\partial P}{\partial t} + \frac{\rho a^2}{A}\frac{\partial Q}{\partial x} = 0. \qquad [2.34]$$

## 2.3 General Remarks on the Momentum and Continuity Equations

The momentum and continuity equations derived above, form a set of partial differential equations describing transient compressible flow in a pipe. A numerical solution method is now needed. Examining the roots of the characteristic equation will determine the type of these equations and offer some guidance as to the best method of solution [15]. Equations [2.11] and [2.34] can be represented in matrix form as:

$$\frac{\partial}{\partial t}\begin{Bmatrix} \mathrm{Q} \\ \mathrm{P} \end{Bmatrix} + \frac{1}{A}\begin{bmatrix} Q & A^2/\rho \\ \rho a^2 & Q \end{bmatrix}\frac{\partial}{\partial x}\begin{Bmatrix} \mathrm{Q} \\ \mathrm{P} \end{Bmatrix} = \begin{Bmatrix} -Ag\dfrac{\partial z}{\partial x} - \dfrac{fQ|Q|}{2DA} \\ 0 \end{Bmatrix}. \qquad [2.35]$$

According to control theory, the eigenvalues, $\lambda$, of the square matrix on the left hand side of equation [2.35] give the roots of the equation set and therefore determine the type of the equation set [44]. The eigenvalues are determined as:

$$B = \begin{bmatrix} Q & A^2/\rho \\ \rho a^2 & Q \end{bmatrix},$$

$$D(\lambda) = \det(B - \lambda I) = \begin{bmatrix} Q-\lambda & A^2/\rho \\ \rho a^2 & Q-\lambda \end{bmatrix} = 0,$$

$$(Q-\lambda)^2 = A^2 a^2,$$

$$\lambda = Q \pm Aa.$$

The eigenvalues, and thus the roots of the characteristic equation, are real and distinct and therefore the equations are a set of non-linear hyperbolic partial differential equations; an equation set that represents wave propagation [15].

## 2.4 The Approximate Equations (Method of Characteristics)

The method of characteristics is based on the momentum and continuity equations derived above. These equations are repeated as:

$$\frac{\partial Q}{\partial t} + \frac{Q}{A}\frac{\partial Q}{\partial x} + \frac{A}{\rho}\frac{\partial}{\partial x}(P + \rho g z) + \frac{fQ|Q|}{2DA} = 0, \qquad [2.11]$$

$$\frac{Q}{A}\frac{\partial P}{\partial x} + \frac{\partial P}{\partial t} + \frac{\rho a^2}{A}\frac{\partial Q}{\partial x} = 0. \qquad [2.34]$$

In order to come up with a numerical technique to solve these two equations, an approximation must be made. The spatial or convective component of the total derivative, $D/Dt$, of both dependent variables, $Q$ and $P$ can be neglected when compared to other terms. When both spatial and time variation terms of $Q$ and $P$ appear in the same

equation, the spatial variation may be neglected since it is much smaller than the time-varying variation [16]. Therefore the momentum and continuity equations may be approximated as:

$$L1 = \frac{\partial Q}{\partial t} + \frac{A}{\rho}\frac{\partial P}{\partial x} + Ag\frac{\partial z}{\partial x} + \frac{fQ|Q|}{2DA} = 0, \qquad [2.36]$$

$$L2 = \frac{A}{\rho}\frac{\partial P}{\partial t} + a^2\frac{\partial Q}{\partial x} = 0. \qquad [2.37]$$

$\partial z/\partial x$ represents the slope of the pipeline and therefore can be written as an ordinary differential equation $dz/dx$. The momentum and continuity equations may be combined in a linear fashion by the use of an unknown multiplier $\Lambda$. Any two real, distinct values of $\Lambda$ will yield two equations equivalent to the momentum and continuity equations. The linear combination is given as:

$$L = \Lambda L1 + L2 = \Lambda\left(\frac{\partial Q}{\partial t} + \frac{A}{\rho}\frac{\partial P}{\partial x} + Ag\frac{\partial z}{\partial x} + \frac{fQ|Q|}{2DA}\right) + \frac{A}{\rho}\frac{\partial P}{\partial t} + a^2\frac{\partial Q}{\partial x} = 0. \qquad [2.38]$$

Grouping terms in equation [2.38] gives:

$$\left(\Lambda\frac{\partial Q}{\partial t} + a^2\frac{\partial Q}{\partial x}\right) + \frac{A}{\rho}\left(\frac{\partial P}{\partial t} + \Lambda\frac{\partial P}{\partial x}\right) + \Lambda Ag\frac{\partial z}{\partial x} + \frac{\Lambda fQ|Q|}{2DA} = 0. \qquad [2.39]$$

Now by choosing two appropriate values for $\Lambda$, [2.39] can be simplified. The two bracketed sections appear to be total derivatives of $Q$ and $P$ with respect to some velocity $dx/dt$. If the first bracketed set of terms is to be replaced by the total derivative multiplied by $\Lambda$, then:

$$\Lambda\frac{dQ}{dt} = \Lambda\left(\frac{\partial Q}{\partial t} + \frac{dx}{dt}\frac{\partial Q}{\partial x}\right) = \Lambda\left(\frac{\partial Q}{\partial t} + \frac{a^2}{\Lambda}\frac{\partial Q}{\partial x}\right),$$

and,

$$\frac{dx}{dt} = \frac{a^2}{\Lambda}. \qquad [2.40]$$

Similarly the second bracketed terms give:

$$\frac{dP}{dt} = \left(\frac{\partial P}{\partial t} + \frac{dx}{dt}\frac{\partial P}{\partial x}\right) = \left(\frac{\partial P}{\partial t} + \Lambda\frac{\partial P}{\partial x}\right),$$

and,

$$\frac{dx}{dt} = \Lambda. \qquad [2.41]$$

Therefore from [2.40] and [2.41], $\Lambda$ can be computed as:

$$\Lambda^2 = a^2 \qquad or \qquad \Lambda = \pm a, \qquad [2.42]$$

and equation [2.39] becomes:

$$a\frac{dQ}{dt} + \frac{A}{\rho}\frac{dP}{dt} + Aag\frac{dz}{dx} + \frac{afQ|Q|}{2DA} = 0. \qquad [2.43]$$

Dividing through by wave speed *a*, gives:

$$\frac{dQ}{dt} + \frac{A}{a\rho}\frac{dP}{dt} + Ag\frac{dz}{dx} + \frac{fQ|Q|}{2DA} = 0. \qquad [2.44]$$

which is valid for:

$$\frac{dx}{dt} = \pm a. \qquad [2.45]$$

Since equation (44) is valid when $dx/dt = \pm a$, it can be written as two separate equations, called the compatibility equations. These two equations are also known as the $C^+$ and $C^-$ equations since they both only exist along so called $C^+$ and $C^-$ characteristic lines. They are given as:

$$\left.\begin{aligned} &\frac{dQ}{dt} + \frac{A}{a\rho}\frac{dP}{dt} + Ag\frac{dz}{dx} + \frac{fQ|Q|}{2DA} = 0, \\ &dx/dt = +a, \end{aligned}\right\} \quad C^+ \qquad [2.46]\ [2.47]$$

$$\left.\begin{aligned} &\frac{dQ}{dt} - \frac{A}{a\rho}\frac{dP}{dt} + Ag\frac{dz}{dx} + \frac{fQ|Q|}{2DA} = 0, \\ &dx/dt = -a. \end{aligned}\right\} \quad C^- \qquad [2.48]\ [2.49]$$

The idea of characteristic lines can be better understood by drawing a characteristic grid. Figure 2.3 displays the characteristic grid. The pipeline shown below the grid is broken into a number of "reaches" of length $\Delta x$. In a real system optimal pipe reach lengths may not be acquired due to the position of important nodes, but for the sake of this discussion all reaches are of length $\Delta x$. Once the shortest reach length is determined, the time step, $\Delta t$, is calculated based on equation [2.47]. Data at each $\Delta x$ position are calculated for every $\Delta t$ interval.

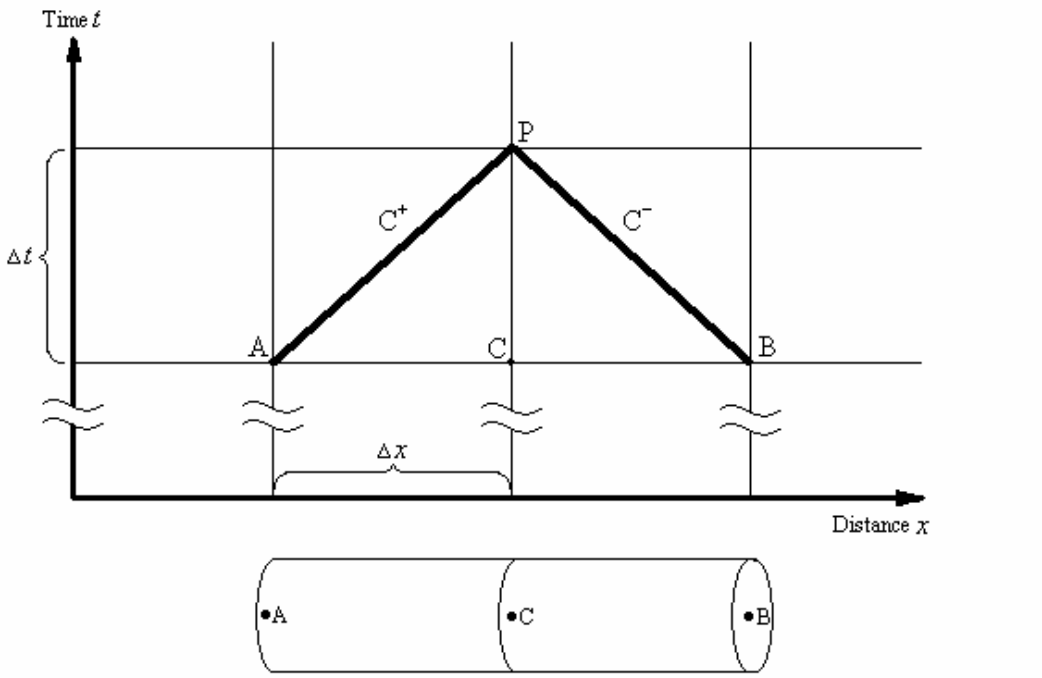


**Figure 2. 3: Method of Characteristics Grid**

Equation [2.47] is satisfied by the $C^+$ line joining point A and D. If pressure and flow are known at A then equation [2.46] can be integrated from A to D. The resulting equation will be in terms of unknown pressure and flow at point D. Similarly [2.49] is satisfied by the $C^-$ line joining points B and D. Knowing pressure and flow at B, equation [2.48] can be integrated along line BD to obtain a second equation relating pressure and flow at point D. Thus at point D, there are two equations and two unknowns ($P_D$ and $Q_D$) Thus pressure and flow at each reach may be calculated throughout time, building a dynamic picture. It is important to note that all $P$'s and $Q$'s must be well defined at $t = 0$ along $x$ to start the integration process. It is now necessary to show how these equations can be integrated along these characteristics.

### 2.4.1 Discretization

If initial conditions are known at A and B, equations [2.46] and [2.48] can be integrated. Multiplying [2.46] by $adt = dx$ and integrating across line AD gives:

$$a\int_{Q_A}^{Q_D}\frac{dQ}{dt}dt+\frac{A}{\rho}\int_{P_A}^{P_D}\frac{dP}{dt}dt+Ag\int_{z_A}^{z_D}\frac{dz}{dx}dx+\frac{f}{2DA}\int_{x_A}^{x_D}Q|Q|dx=0\,. \qquad [2.50]$$

The integration is straightforward with the exception of the last term which models the change in flow with respect to distance along the pipe $\Delta x$. The spatial variation of $Q$ is unknown and therefore an approximation of the last term in [2.50] is needed. A first order approximation is satisfactory in cases where the friction term does not dominate [15]. Therefore, the last term of equation [2.50] may be approximated as $Q_A|Q_A|\Delta x$. Integrating equation [2.50] gives:

$$a(Q_D-Q_A)+\frac{A}{\rho}(P_D-P_A)+Ag(z_D-z_A)+\frac{f\Delta x}{2DA}Q_A|Q_A|=0\,, \qquad [2.52]$$

and solving for $P_D$ gives:

$$C^+:\;\; P_D=P_A-\frac{a\rho}{A}(Q_D-Q_A)-\rho g(z_D-z_A)-\frac{f\rho\Delta x}{2DA^2}Q_A|Q_A|\,, \qquad [2.53]$$

Similarily, integrating along the $C^-$ characteristic line gives:

$$C^-:\;\; P_D=P_B+\frac{a\rho}{A}(Q_D-Q_B)+\rho g(z_D-z_B)+\frac{f\rho\Delta x}{2DA^2}Q_B|Q_B|\,, \qquad [2.54]$$

Pressure is replaced by hydraulic head to simplify the model. Hydraulic head is given by:

$$H=\frac{P}{\rho g}+z\,.$$

Subscript notation will now be introduced to clearly define variables. The subscripts $i$ and $j$ shown as $Q_{ij}$ denote the spatial and temporal locations of measurement D.

Therefore, replacing pressure with head, introducing the new notation and simplifying gives:

$$C^+ : \; H_{ij} = C_P - BQ_{ij}\,, \tag{2.55}$$

$$C^- : \; H_{ij} = C_M + BQ_{ij}\,. \tag{2.56}$$

in which the constants are given as:

$$C_P = H_{i-1,j-1} + Q_{i-1,j-1}\left[B - R\left|Q_{i-1,j-1}\right|\right], \tag{2.57}$$

$$C_M = H_{i+1,j-1} - Q_{i+1,j-1}\left[B - R\left|Q_{i+1,j-1}\right|\right], \tag{2.58}$$

$$B = \frac{a}{Ag}\,, \tag{2.59}$$

$$R = \frac{f\Delta x}{2gDA^2}\,. \tag{2.60}$$

Equations [2.55] to [2.60] can be easily implemented into computer code to solve for head and flow throughout a pipeline. Initial conditions, which are needed to start the solution process, are calculated using a steady-state solution. At pipe inlets the $C^+$ equation does not exist, since no pipe exists upstream of the inlet, and therefore a supplementary boundary equation must be found. Similarly, at outlets the $C^-$ equation does not exist, since no pipe exists downstream of the outlet, and therefore another boundary equation is needed. Typical boundary conditions include pumps, reservoirs and valves. Once the boundary equations are known, a complete solution to the transient pipe problem exists.

# Chapter 3: The Extended Kalman Filter

## 3.1 Introduction

This chapter presents the theory associated with the Extended Kalman Filter (EKF). The EKF is used in this thesis for parameter estimation and serves as a tool for estimating pipeline leakage, given a model of the pipeline and a record of system inputs and outputs. The EKF is an extension of the Kalman Filter (KF). It has been widely used for the estimation of states and parameters within non-linear models. In presenting the theory of the EKF, it is first necessary to discuss the basic Kalman Filter.

The Kalman Filter is named after Rudolph E. Kalman, who published a paper in 1960 describing a recursive solution to the "discrete-data linear filtering problem" [45]. Its extensive use over the past 40 years, within a diverse array of fields, demonstrates that the Kalman Filter is the most useful state estimation tool to emerge from the state variable approach of modern control theory [46]. A few of the areas in which the filter has been used are navigation, tracking and guidance, condition monitoring, water and air quality, and speech recognition and enhancement. The filter is a recursive state estimation tool that provides a robust, linear, unbiased solution of the least-squares method. In order to understand the theory of the KF, the reader should understand the principles of state space modeling and probability statistics. Appendix A presents the theory behind state space modeling. Appendix B gives the relevant background of probability statistics. Readers unfamiliar with these topics are encouraged to first read these appendices before continuing on with this chapter.

## 3.2 Kalman Filtering

The Kalman Filter is a recursive predictor-corrector algorithm for processing discrete measurements (inputs), with the aid of a system model, into optimal state estimates (outputs). The filter minimizes the estimated error covariance in a linear stochastic system. Stochastic systems are those that have an associated random portion or

corruption due to noise. There are two types of noise, namely process noise and measurement noise. The first is the noise associated with the system and its states. Measurement noise is noise that comes from sensors and the instrumentation used in monitoring the process. The Kalman Filter is capable of handling systems where there is a high degree of uncertainty in the data; therefore it is a prime candidate for pipeline leak detection.

### 3.2.1 Discrete State Space Model

State space models are simply a convenient representation of the governing equations, that make tractable what would otherwise be a notational nightmare [47]. In a stochastic state space model, the notation is such that current states are only functions of prior states, inputs and random noise. The state and measurement equations are given as:

$$x_{k+1} = \phi_k x_k + G_k u_k + w_k, \qquad [3.1]$$

$$z_{k+1} = H_k x_{k+1} + v_{k+1}, \qquad [3.2]$$

where, $x_k = (n \times 1)$ state vector at time $t_k$,

$\phi_k = (n \times n)$ system or transition matrix of constants for time $t_k$,

$G_k = (n \times r)$ input matrix,

$u_k = (r \times 1)$ input vector

$w_k = (n \times 1)$ vector of random system disturbances characterized by zero mean white noise,

$z_{k+1} = (p \times 1)$ vector of defined measurements,

$H_k = (p \times n)$ output matrix that linearly connects outputs and states, and

$v_{k+1} = (p \times 1)$ vector of white measurement noise.

The process and measurement noise, $w_k$ and $v_k$, repectively, are assumed white. The term white describes a signal that contains an equal distribution of all frequencies, akin to white light that has the property of containing an equal distribution of all frequencies, or

colours, of light. White noise is a random, zero mean sequence that is uncorrelated temporally. However, the individual elements of the noise vectors have known correlations, at any point in time [48]. These correlations are denoted by $Q_k$ and $R_k$.

$$E[w_k] = 0 \qquad E[v_k] = 0, \tag{3.3}$$

$$\mathrm{cov}[w_k, w_j] = E[w_k, w_j^T] = \begin{cases} Q_k, & j = k \\ 0, & j \neq k \end{cases}, \tag{3.4}$$

$$\mathrm{cov}[v_k, v_j] = E[v_k, v_j^T] = \begin{cases} R_k, & j = k \\ 0, & j \neq k \end{cases}, \tag{3.5}$$

where, $E[\ ]$ denotes the expectation, a zero mean process, and $\mathrm{cov}[\ ]$ denotes the covariance.

### 3.2.2 The Filtering Process

The Kalman Filter works by blending two inputs, the system measurements and their respective model predictions, with a gain factor, denoted $K_k$. The noise corrupted measurements $z_k$ offer knowledge from the system that is not known by the assumed model. The gain factor offers a way of "educating" the model by incorporating knowledge from measurements through a comparison of the states that can be measured and their equivalent predicted states. The predictor-corrector algorithm, implemented by the filter, is explained in this section.

The predictor phase, also called the time update phase, uses the system model to determine an "unrefined" estimate of the states, at time *t*, based on a prior "refined" estimates or initial estimates, at time *t*-1. The unrefined estimate, determined by the predictor phase, is denoted $\hat{x}_k^-$. In general, unrefined estimates are denoted by a "minus sign" superscript. The refined estimate, determined by the corrector phase, is denoted $\hat{x}_k$. The process can not be started without an initial estimate of the states, an input vector, and their associated error covariance matrices. In the work of this thesis, a steady state solution of the pipeline is determined for the initial estimates. However, the Kalman

Filter may be started by simply assigning zero values to the states and the filter will converge due to its nature.

An unrefined estimate of covariance, $P_k^-$, is also determined in the predictor phase of the filter. This covariance matrix will be discussed in the next section dealing with the computational origins of the filter. The predictor equations are given as:

$$\hat{x}_{k+1}^- = \phi_k \hat{x}_k + G_k u_k, \quad [3.6]$$

$$P_{k+1}^- = \phi_k P_k \phi_k^T + Q_k. \quad [3.7]$$

The corrector phase, also called the measurement update phase, starts by first calculating the corrective gain, $K_k$, from the unrefined estimate of covariance. The error between measurement and unrefined states is then multiplied by the corrective gain and added to the unrefined estimate to obtain the refined estimate. Finally the refined covariance estimate is determined with the help of the corrective gain. Again, the corrective gain and refined covariance equations will be discussed further in subsequent sections. The corrective phase equations are given as:

$$K_k = P_k^- H_k^T \left( H_k P_k^- H_k^T + R_k \right)^{-1}, \quad [3.8]$$

$$\hat{x}_k = \hat{x}_k^- + K_k \left( z_k - H_k \hat{x}_k^- \right), \quad [3.9]$$

$$P_k = \left( I - K_k H_k \right) P_k^-. \quad [3.10]$$

As mentioned previously, the Kalman gain, $K_k$, blends measurement information and system inputs. The covariance matrices, $R_k$ and $Q_k$ act as weighting factors for measurement and input data in the determination of $K_k$. This can be seen in equations [3.7] and [3.8] and by illustrating a simple example. Consider a system where measurement noise is very small in comparison to input noise. Thus $R_k << Q_k$. If $R_k$ is assumed negligible, then from equation [3.8] the Kalman gain is given as $K_k = H_k^{-1}$.

Substituting this gain into equation [3.9] gives $\hat{x}_k = z_k H_k^{-1}$. Therefore the state predictions are only functions of the measurement data and plant inputs are neglected. Likewise, if model predictions are considered highly accurate and measurements considered noisy, the filter will weight the model predictions more and the measurements with much less degree [49].

Equations [3.6] to [3.10] are a recursive process that may be summarized visually by Figure 3.1 below [48].

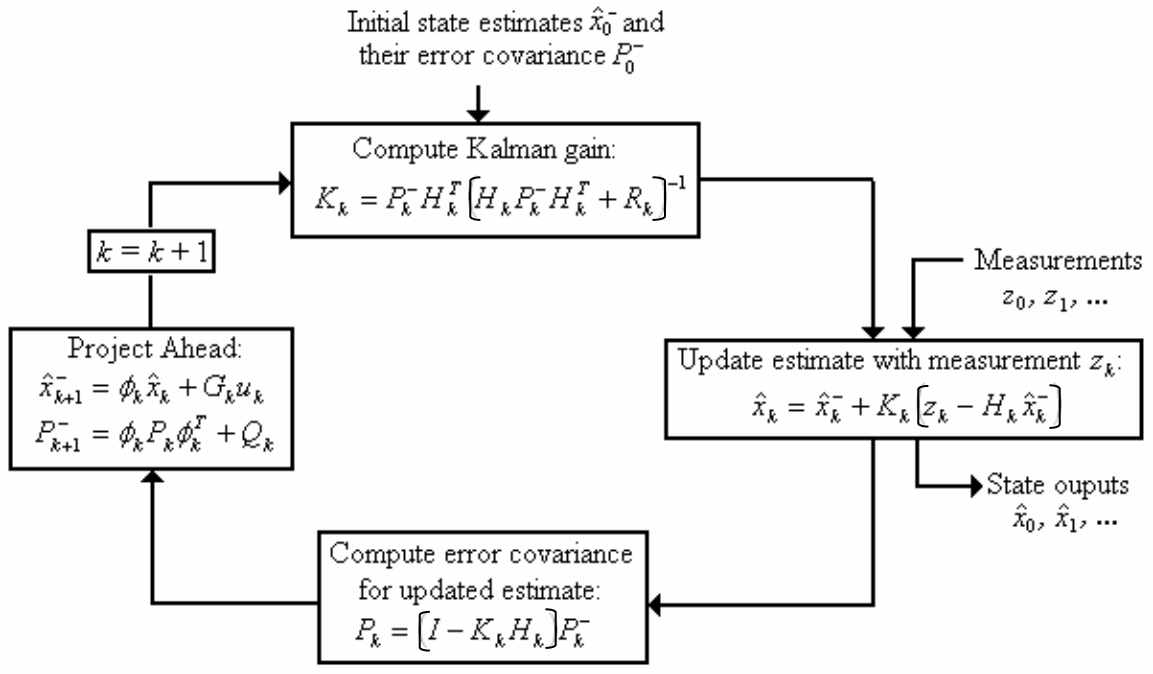


**Figure 3. 1: Kalman Filter Loop**

### 3.2.3 Computational Origins of the Filter

The basis of the filter, as stated before, is the minimization of the error covariance. There are two estimate errors, modeling error and measurement error, that are defined as *a priori* and *a posteriori* [47]. *A priori* estimate errors are given as the difference between actual states and those unrefined states predicted by the model. *A posteriori* errors are

those between actual states and states predicted with knowledge from the system, i.e. measurements (refined). These estimate errors are given as:

$$e_k^- \equiv x_k - \hat{x}_k^-, \ (a\ priori) \qquad [3.11]$$

$$e_k \equiv x_k - \hat{x}_k . \ (a\ posteriori) \qquad [3.12]$$

The *a priori* and *a posteriori* error covariances are given as:

$$P_k^- = E\left[e_k^- e_k^{-T}\right] = E\begin{bmatrix} (e_{1k}^-)^2 & e_{1k}^- e_{2k}^- & \cdots & \\ e_{2k}^- e_{1k}^- & (e_{2k}^-)^2 & & \\ \vdots & & \ddots & \\ & & & (e_{nk}^-)^2 \end{bmatrix} = E\left[\left(x_k - \hat{x}_k^-\right)\left(x_k - \hat{x}_k^-\right)^T\right], \qquad [3.13]$$

$$P_k = E\left[e_k e_k^T\right] = E\begin{bmatrix} (e_{1k})^2 & e_{1k} e_{2k} & \cdots & \\ e_{2k} e_{1k} & (e_{2k})^2 & & \\ \vdots & & \ddots & \\ & & & (e_{nk})^2 \end{bmatrix} = E\left[\left(x_k - \hat{x}_k\right)\left(x_k - \hat{x}_k\right)^T\right]. \qquad [3.14]$$

Note that in equations [3.13] and [3.14] the diagonal matrix terms represent the mean squared error or the variance of $e_{ik}$. The final goal is to find an expression for the Kalman gain $K_k$ that minimizes these mean squared error values. Therefore an expression relating $K_k$ and the error covariance must first be found. The derivative of equation [3.14] with respect to $K_k$ gives the optimal least squares result for the Kalman gain. The expression relating $K_k$ and $P_k$ is found by combining equations [3.14] and [3.9] to obtain:

$$P_k = E\left[\left(x_k - \hat{x}_k^- - K_k\left(z_k - H_k \hat{x}_k^-\right)\right)\left(x_k - \hat{x}_k^- - K_k\left(z_k - H_k \hat{x}_k^-\right)\right)^T\right]. \qquad [3.15]$$

Substituting in equations [3.2] and [3.11] gives:

$$P_k = E\left[\left(e_k^- - K_k\left(H_k x_k + v_k - H_k \hat{x}_k^-\right)\right)\left(e_k^- - K_k\left(H_k x_k + v_k - H_k \hat{x}_k^-\right)\right)^T\right]. \qquad [3.16]$$

Simplifying the first bracketed expression of equation [3.16]:

$$\begin{aligned} & e_k^- - K_k\left(H_k x_k + v_k - H_k \hat{x}_k^-\right) \\ &= e_k^- - K_k H_k x_k - K_k v_k + K_k H_k \hat{x}_k^- \\ &= e_k^- - K_k H_k\left(x_k - \hat{x}_k^-\right) - K_k v_k \\ &= e_k^- - K_k H_k e_k^- - K_k v_k \\ &= \left[I - K_k H_k\right] e_k^- - K_k v_k . \end{aligned} \qquad [3.17]$$

The second expression of equation [3.16] may be manipulated to:

$$\begin{aligned} & \left[e_k^- - K_k\left(H_k x_k + v_k - H_k \hat{x}_k^-\right)\right]^T \\ &= e_k^{-T} - \left(H_k x_k + v_k - H_k \hat{x}_k^-\right)^T K_k{}^T \\ &= e_k^{-T} - \left(H_k e_k^- + v_k\right)^T K_k{}^T \\ &= e_k^{-T} - \left(H_k e_k^- + v_k\right)^T K_k{}^T \\ &= e_k^{-T} - \left(e_k^{-T} H_k{}^T + v_k{}^T\right) K_k{}^T \\ &= e_k^{-T}\left(I - H_k{}^T K_k{}^T\right) - v_k{}^T K_k{}^T . \end{aligned} \qquad [3.18]$$

Combining [3.17] and [3.18] to obtain the expression of the error covariance matrix gives:

$$\begin{aligned} P_k &= E\left[\left[\left[I - K_k H_k\right] e_k^- - K_k v_k\right]\left[e_k^{-T}\left(I - H_k{}^T K_k{}^T\right) - v_k{}^T K_k{}^T\right]\right] \\ &= E\left[\left[I - K_k H_k\right] e_k^- e_k^{-T}\left[I - H_k{}^T K_k{}^T\right] + K_k v_k v_k{}^T K_k{}^T\right] \\ P_k &= \left[I - K_k H_k\right] E\left[e_k^- e_k^{-T}\right]\left[I - H_k{}^T K_k{}^T\right] + K_k E\left[v_k v_k{}^T\right] K_k{}^T . \end{aligned} \qquad [3.19]$$

Remember that from equation [3.13], $P_k^- = E\left[e_k^- e_k^{-T}\right]$ is the *a priori* error covariance. Also, from equation [3.5], $R_k = E\left[v_k v_k^T\right]$ is the covariance matrix of the measurement noise. Rewriting $P_k$ gives:

$$P_k = \left[I - K_k H_k\right] P_k^- \left[I - H_k{}^T K_k{}^T\right] + K_k R_k K_k{}^T . \qquad [3.20]$$

Equation [3.20] is the general expression for the error covariance and is valid for any value of $K_k$. As stated before, the goal of the Kalman Filter is to determine a minimized mean square error solution. Since the diagonal or trace of $P_k$ represents the mean square error, setting the derivative of the trace of $P_k$ equal to zero will result in the optimal gain $K_k$.

To determine the derivative of equation [3.20], two matrix differentiation procedures are needed. These are given as:

$$\frac{d\left[trace(xy)\right]}{dx} = y^T \quad (xy \text{ must be square}), \qquad [3.21]$$

$$\frac{d\left[trace\left(x^T A x\right)\right]}{dx} = 2Ax \quad (\text{if } A \text{ is symmetric}). \qquad [3.22]$$

It is also useful to identify that, since $P_k^-$ equals $P_k^{-T}$, the trace of $P_k^- H_k{}^T K_k{}^T$ is equal to the trace of the transpose $K_k H_k P_k^-$. Expanding equation [3.20] and considering the above mentioned comment gives:

$$P_k = P_k^- - 2K_k H_k P_k^- + K_k \left[H_k P_k^- H_k{}^T + R_k\right] K_k{}^T . \qquad [3.23]$$

Differentiating the trace of equation [3.23], considering [3.21] and [3.22] gives:

$$\frac{d}{dK_k}\left(trace(P_k)\right) = -2P_k^- H_k{}^T + 2\left[H_k P_k^- H_k{}^T + R_k\right]K_k . \qquad [3.24]$$

Setting this derivative equal to zero and solving for the gain $K_k$ gives:

$$K_k = \frac{P_k^- H_k{}^T}{H_k P_k^- H_k{}^T + R_k} . \qquad [3.25]$$

Equation [3.25] gives the optimal gain value. Note that the optimal gain is a function of *a priori* error covariance and not *a posteriori*. The covariance matrix that is associated with this gain may now be determined by substituting equation [3.25] into equation [3.23] as:

$$\begin{aligned} P_k &= P_k^- - 2\frac{P_k^- H_k{}^T}{H_k P_k^- H_k{}^T + R_k} H_k P_k^- + P_k^- H_k{}^T \frac{H_k P_k^-}{H_k P_k^- H_k{}^T + R_k}, \\ &= P_k^- - P_k^- H_k{}^T \left(H_k P_k^- H_k{}^T + R_k\right)^{-1} H_k P_k^- , \\ &= P_k^- - K_k H_k P_k^- , \\ P_k &= \left(I - K_k H_k\right)P_k^- . \end{aligned} \qquad [3.26]$$

Equation [3.26] is the same as [3.10]. This covariance matrix is valid only for the optimal gain determined by equation [3.25]. All but the *a priori* estimate of the covariance matrix have now been determined. The *a priori* covariance matrix is derived by substituting equations [3.1] and [3.6] into the definition of the *a priori* error. This is given as:

$$e_{k+1}^- \equiv x_{k+1} - \hat{x}_{k+1}^- , \ (a\ priori) \qquad [3.11]$$

$$\begin{aligned} e_{k+1}^- &\equiv \left(\phi_k x_k + G_k u_k + w_k\right) - \left(\phi_k x_k - G_k u_k\right), \\ &\equiv \left(\phi_k x_k + w_k\right) - \phi_k \hat{x}_k , \\ e_{k+1}^- &\equiv \phi_k e_k + w_k . \end{aligned} \qquad [3.27]$$

Note that there is no correlation between $e_k$ and $w_k$ because $w_k$ is a process noise for the next time step [48]. Therefore the *a priori* covariance matrix is given as:

$$\begin{aligned} P_{k+1}^{-} &= E\left[e_{k+1}^{-} e_{k+1}^{-T}\right] = E\left[\left(\phi_k e_k + w_k\right)\left(\phi_k e_k + w_k\right)^T\right], \\ &= E\left[\left(\phi_k e_k + w_k\right)\left(e_k{}^T \phi_k{}^T + w_k{}^T\right)\right], \\ &= E\left[\phi_k e_k e_k{}^T \phi_k{}^T + \phi_k e_k w_k{}^T + w_k e_k{}^T \phi_k{}^T + w_k w_k{}^T\right]. \end{aligned}$$

Now, remember that from equation [3.4], $Q_k = E\left[w_k w_k{}^T\right]$, and from equation [3.14], that $P_k = E\left[e_k e_k^T\right]$. Equation [3.3] states that the process noise is non biased, or $E\left[w_k\right] = 0$. Therefore the *a priori* estimate of the covariance matrix may be simplified to:

$$P_{k+1}^{-} = \phi_k P_k \phi_k{}^T + Q_k . \qquad [3.28]$$

The five Kalman Filter equations have now been derived. These equations form a recursive loop of prediction and correction. The equations may be summarized as follows:

Prediction Equations:

$$\hat{x}_{k+1}^{-} = \phi_k \hat{x}_k + G_k u_k , \qquad [3.6]$$

$$P_{k+1}^{-} = \phi_k P_k \phi_k^T + Q_k . \qquad [3.28]$$

Corrector Equations:

$$K_k = P_k^{-} H_k^T \left(H_k P_k^{-} H_k^T + R_k\right)^{-1}, \qquad [3.25]$$

$$\hat{x}_k = \hat{x}_k^{-} + K_k\left(z_k - H_k \hat{x}_k^{-}\right), \qquad [3.9]$$

$$P_k = \left(I - K_k H_k\right) P_k^{-} . \qquad [3.26]$$

Determining the proper measurement and system noise covariance matrices often relies on experience and tuning. Often a trial and error approach, for determining $R_k$ and $Q_k$,

is used in order to effectively tune the filter and avoid divergence. As indicated in the literature, there does not exist one single cure for all numerical problems and the choice of $R_k$ and $Q_k$, in some instances, may be very arbitrary [48]. The measurement noise covariance is a measure of sensor accuracy. Therefore, offline tests may be preformed to determine the statistical variance of the measurement device and therefore $R_k$.

Obtaining a value of the system noise is often much more difficult since the actual system may not be physically observable or concrete. The system noise covariance may therefore be thought of as the uncertainty of the process model. In cases where the process model may be overly simplistic or inaccurate, good results may still be obtained by simply adding enough uncertainty to the system noise covariance, $Q_k$. In this way, the filter will put more merit in measurements and less in the model predictions.

## 3.3 Extended Kalman Filter

In cases where the describing dynamics of the system are non-linear, the Extended Kalman Filter EKF must be implemented. Such is the case in the work of this thesis. The Extended Kalman Filter works in the same way as the ordinary Kalman Filter. It differs from the ordinary Kalman Filter because of the need to linearize the state equations around the most recent state estimate for each time interval. The linearization is done using a Taylor series approximation.

The Taylor series linearization is done about the most recent approximation of the state vector. Therefore, the measurement equation is linearized about the *a priori* state estimate since this is the most recent estimate of states, and the state equation is linearized about the *a posteriori* state estimate.

The non-linear model may be described by the following state equations:

$$x_{k+1} = f\left[x_k\right] + u_k + w_k, \quad \text{(State equation)} \qquad [3.29]$$

$$z_k = h[x_k] + v_k, \quad \text{(Measurement equation)} \tag{3.30}$$

where $f[x_k]$ and $h[x_k]$ may both be nonlinear functions. The algorithm starts, at step *k*, with an estimate of the state, $\hat{x}_k$, and its covariance $P_k$. The first step filter step is to calculate an *a priori* estimate of the state vector. This is given as:

$$\hat{x}_{k+1}^{-} = f[\hat{x}_k] + u_k. \tag{3.31}$$

in which $f[\hat{x}_k]$ is the solution of the nonlinear function about the most current state estimate $\hat{x}_k$.

The next step is to calculate the covariance matrix of this preliminary state estimate. In order to do this, equation [3.31] must first be linearized about its current state $\hat{x}_k$. In general, a nonlinear function, $f[x_k]$, may be linearized about $\hat{x}_k$ by a Taylor series approximation as:

$$f[x_k] \approx f[\hat{x}_k] + J_x[x_k - \hat{x}_k], \tag{3.32}$$

in which $J_x$ is the Jacobian of $f[x_k]$ evaluated at $\hat{x}_k$. The Jacobian operation is given as:

$$J_x = \left.\frac{\partial f[x_k]}{\partial x}\right|_{\hat{x}_k} = \begin{bmatrix} \frac{\partial f[x_{1k}]}{\partial x_1} & \frac{\partial f[x_{1k}]}{\partial x_2} & \cdots \\ \frac{\partial f[x_{2k}]}{\partial x_1} & \frac{\partial f[x_{2k}]}{\partial x_2} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}_{\hat{x}_k}. \tag{3.33}$$

The *a priori* covariance matrix may now be computed in a similar fashion to equation [3.28]. The difference being that the Extended Kalman Filter uses the Jacobian of the nonlinear state relation, while the Kalman Filter uses the linear transition matrix in the calculation of $P_{k+1}^{-}$. The *a priori* covariance matrix is therefore given as:

$$P_{k+1}^{-} = J_x P_k J_x^T + Q_k . \qquad [3.34]$$

System measurements may now be used to form *a posteriori* or informed state and covariance estimates. It should be noted however that in a case where no measurements are available, equations [3.31] and [3.34] immediately become the final or *a posteriori* estimates. In the case where measurements are available equation [3.30] must be linearized. This linearization is done about the *a priori* estimate, $\hat{x}_{k+1}^{-}$, since it is the most recent estimate of states. In general, a nonlinear function, $h[x_k]$, may be linearized about $\hat{x}_{k+1}^{-}$ by a Taylor series approximation as:

$$h[x_k] \approx h[\hat{x}_{k+1}^{-}] + J_h[x_k - \hat{x}_{k+1}^{-}]. \qquad [3.35]$$

in which $J_h$ is the Jacobian of $h[x_k]$ evaluated at $\hat{x}_{k+1}^{-}$. The Jacobian operation is given as:

$$J_h = \frac{\partial h[x_k]}{\partial x}\bigg|_{\hat{x}_{k+1}^{-}} = \begin{bmatrix} \frac{\partial h[x_{1k}]}{\partial x_1} & \frac{\partial h[x_{1k}]}{\partial x_2} & \cdots \\ \frac{\partial h[x_{2k}]}{\partial x_1} & \frac{\partial h[x_{2k}]}{\partial x_2} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}_{\hat{x}_{k+1}^{-}} . \qquad [3.36]$$

The measurement Jacobian may now be used to compute the Kalman gain $K_k$. This operation is done in a similar fashion to equation [3.25] as:

$$K_k = P_k^- J_h^T \left( J_h P_k^- J_h^T + R_k \right)^{-1}, \qquad [3.37]$$

where again, the linear matrices of the Kalman Filter are replace by the Jacobians of their non-linear counterparts. The *a posteriori* state matrix may now be determined by:

$$\hat{x}_k = \hat{x}_k^- + K_k \left( z_k - h\left[ \hat{x}_k^- \right] \right). \qquad [3.38]$$

The last step of the algorithm is to compute the *a posteriori* covariance estimate as:

$$P_k = \left( I - K_k J_h \right) P_k^- .$$

The EKF algorithm is represented below in Figure 3.2.

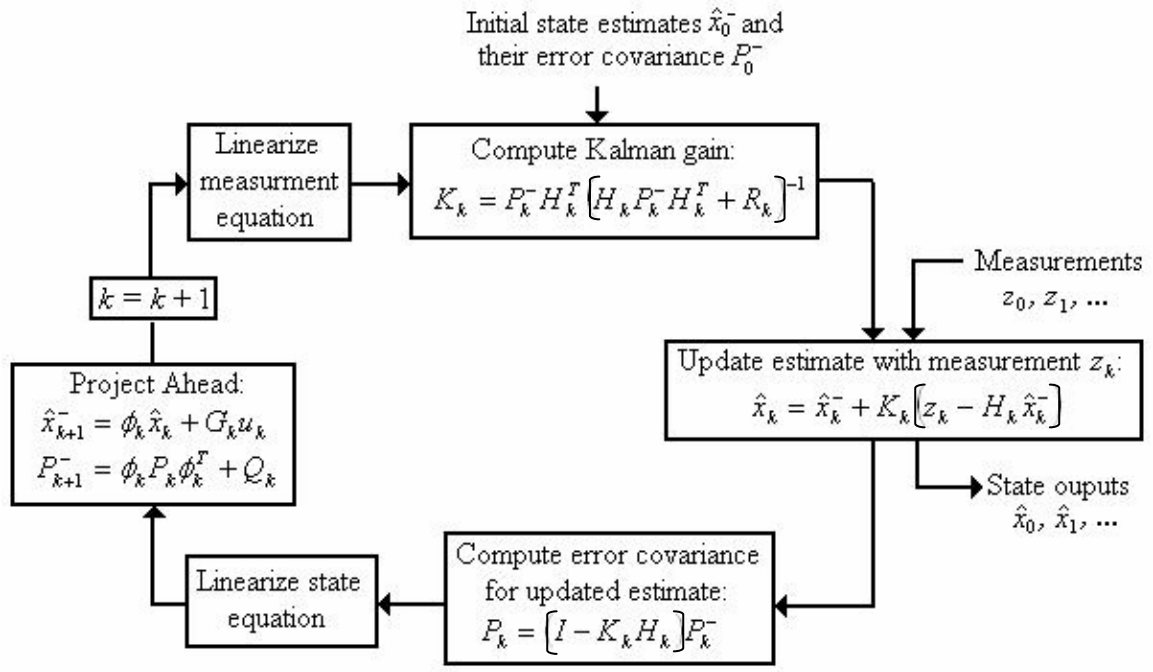


**Figure 3. 2: Extended Kalman Filter Loop**

# Chapter 4: Water Transmission Line Model

## 4.1 Preliminary Remarks

This chapter presents the water transmission line model equation set. The results of the proposed model are compared with the results of a commercial software package, TransAM, and the differences are discussed [50]. The developed equation set is transformed into a state space format, which is required for the implementation of Extended Kalman Filtering.

## 4.2 System Configuration and Equations

The model developed below represents a simple transmission line or a portion of a large distribution system. This model is based roughly on a portion of the SaskWater Saskatoon East raw water transmission line. The model consists of a constant head reservoir of 40 [m] connected to a 0.5 [m] radius, 600 [m] long pipe with a downstream reservoir of 30 [m]. For simplicity, no elevation changes occur within the model. A valve is located just upstream of the 30 [m] reservoir. The model is simplified in order to more clearly identify the effectiveness of the proposed loss detection algorithm. The model configuration is shown below in Figure 4.1.

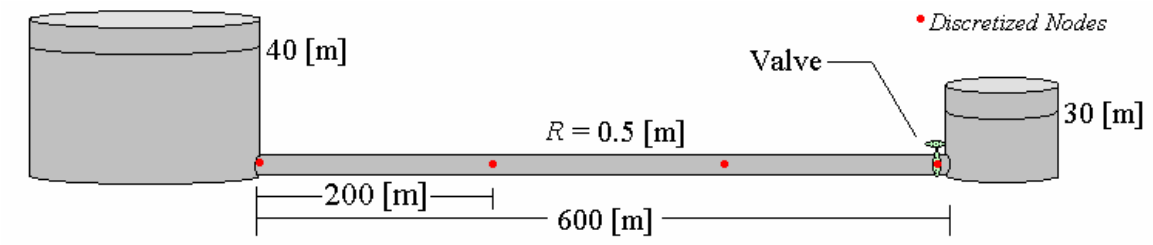


**Figure 4. 1: Two Reservoir Supply Line**

The pipe is segmented into three 200 [m] long sections. Therefore there are four nodes; one node at the upstream or supply reservoir; two interior nodes at 200 [m] and 400 [m] from the supply; and one end node located at the valve and downstream reservoir.

Leakage is assumed to occur at each interior node. The model assumes that there is always zero leakage at the pipe boundary nodes $Q_{L1} = Q_{L4} = 0$. The two interior leaks, $Q_{L2}$ and $Q_{L3}$, are unknown "fictitious" states. These two states may be used to determine the magnitude and location of one leak through linear interpolation.

Notation for flow throughout the pipeline takes the following form, $Q_{21,k}$. The first two-digit subscript represents the spatial position of the state and the second subscript, *k*, represents the temporal position of the state. In this case, the flow is in the beginning of the second pipe segment, at the current time step *k*. Assuming that the head just upstream and downstream of a node is equal, a simpler notation is adopted for pressure head terms. The head a node 2 is simply denoted as $H_{2,k}$. This notation is shown in Figure 4.2.

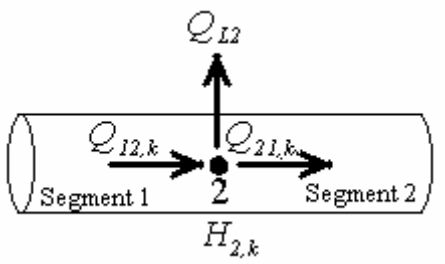


**Figure 4. 2: Subscript Notation**

### 4.2.1 The Supply Reservoir

The cross sectional area of the supply reservoir is large and therefore it is assumed that the reservoir water level remains constant during transient conditions. The energy equation is used to specify the conditions which occur at the reservoir boundary. By combining the energy equation and the $C^-$ equation from the method of characteristics, given in Chapter 2 by equation [2.56], a solution for all system states at the supply reservoir may be found. These two equations are given as:

$$H_{1,k} = H_{R1} - (1+K)\frac{{Q_{11,k}}^2}{2gA^2}, \qquad [4.1]$$

$$Q_{11,k} = \frac{\left(H_{1,k} - H_{2,k-1} + BQ_{12,k-1} - RQ_{12,k-1}\left|Q_{12,k-1}\right|\right)}{B}. \qquad [4.2]$$

$H_{R1}$ represents the supply reservoir head, and $K = 0.5$ is the minor losses at a blunt exit. From Chapter 2, $B = a/gA$ and $R = f\Delta x/\left(2gDA^2\right)$. Substituting for $H_{1,k}$ in equation [4.2] and neglecting the negative sign with the radical term gives:

$$Q_{11,k} = \frac{-B + \sqrt{B^2 - 4\left(\frac{1+K}{2gA^2}\right)\left(H_{2,k-1} - BQ_{12,k-1} + RQ_{12,k-1}\left|Q_{12,k-1}\right| - H_{R1}\right)}}{1+K/gA^2}. \qquad [4.3]$$

$H_{1,k}$ may now be solved using equation [4.1].

### 4.2.2 The Downstream Reservoir and Valve

The downstream reservoir-valve boundary condition is modeled using the valve equation. The development of the valve equation is taken from Chaudhry and Wylie/Streeter [15] [14]. The orifice equation for steady state flow through a valve is given as:

$$Q_{32,0} = (C_d A)_0 \sqrt{2g(H_{4,0} - H_{R2})}, \qquad [4.4]$$

where $Q_{32,0}$ is the steady state flow, $(H_{4,0} - H_{R2})$ is the steady state head loss across the valve, $H_{R2}$ is the downstream reservoir head and $(C_d A)_0$ is the steady state discharge coefficient times the valve opening area. It is assumed that flow is turbulent through the valve as the expected Reynolds numbers is 637,000. All other flows through the valve may be represented as:

$$Q_{32,k} = (C_d A)\sqrt{2g(H_{4,k} - H_{R2})}. \qquad [4.5]$$

Dividing the two orifice equations gives:

$$Q_{32,k} = \frac{Q_{32,0}}{\sqrt{(H_{4,0} - H_{R2})}} \tau \sqrt{(H_{4,k} - H_{R2})}\,, \qquad [4.6]$$

where $\tau$ is the dimensionless valve opening given as:

$$\tau = \frac{(C_d A)}{(C_d A)_0}. \qquad [4.7]$$

Substituting for $H_{4,k}$ using the $C^+$ equation from the Method of Characteristics, equation [2.55], gives:

$$Q_{32,k}{}^2 + C_v B Q_{32,k} - C_v \left( H_{3,k-1} + B Q_{31,k-1} - R \Delta t Q_{31,k-1} \left| Q_{31,k-1} \right| - H_{R2} \right) = 0\,, \qquad [4.8]$$

where $C_v$ is a valve opening coefficient given by:

$$C_v = \frac{(Q_{32,0}\tau)^2}{(H_{4,0} - H_{R2})}. \qquad [4.9]$$

Finally, solving for $Q_{32,k}$ and neglecting the negative sign with the radical term:

$$Q_{32,k} = \frac{-C_v B + \sqrt{(C_v B)^2 + 4C_v \left( H_{3,k-1} + Q_{31,k-1} \left( B - R \left| Q_{31,k-1} \right| \right) - H_{R2} \right)}}{2}. \qquad [4.10]$$

For a simple valve closure, from fully open to fully closed, the dimensionless valve opening is given by the equation set:

$$\tau = 1 \qquad t < t_s ,$$

$$\tau = \left(1 - \frac{t - t_s}{t_c}\right)^{3/2} \qquad t_s \le t \le (t_s + t_c), \qquad [4.11]$$

$$\tau = 0 \qquad t > (t_s + t_c)$$

where $t$ is time, $t_s$ is closure start time, and $t_c$ is the time duration of valve closure. Initially the valve is fully open and therefore $\tau = 1$. Once the valve has fully closed, $\tau = 0$.

The hydraulic head at the valve may now be determined from the $C^+$ equation of the Method of Characteristics as:

$$H_{4,k} = H_{3,k-1} + BQ_{31,k-1} - RQ_{31,k-1}\left|Q_{31,k-1}\right| - BQ_{32,k} . \qquad [4.12]$$

### 4.2.3 Inner Nodes with Leakage

States at the inner nodes may be determined using the $C^+$ and $C^-$ equations for incoming and outgoing flow and an orifice equation for leakage flow. The head corresponding to each of these flows is assumed equivalent. Therefore, for node 2:

$$Q_{12,k} = \left(H_{2,k} + H_{1,k-1} + BQ_{11,k-1} - RQ_{11,k-1}\left|Q_{11,k-1}\right|\right)\Big/ B , \qquad [4.13]$$

$$-Q_{21,k} = \left(-H_{2,k} + H_{3,k-1} - BQ_{22,k-1} + RQ_{22,k-1}\left|Q_{22,k-1}\right|\right)\Big/ B , \qquad [4.14]$$

$$-Q_{L2} = -\lambda_2 \sqrt{H_{2,k}} , \qquad [4.15]$$

where $\lambda_2$ is the unknown leakage area constant for node 2. A summation of flow gives:

$$\sum Q = 0 = \frac{-2H_{2,k} + H_{1,k-1} + BQ_{11,k-1} - RQ_{11,k-1}\left|Q_{11,k-1}\right| + H_{3,k-1} - BQ_{22,k-1} + RQ_{22,k-1}\left|Q_{22,k-1}\right|}{B} - \lambda_2\sqrt{H_{2,k}}$$

[4.16]

Solving for $H_{2,k}$ and neglecting the positive sign with the radical term gives:

$$H_{2,k} = \frac{1}{8}B^2{\lambda_2}^2 + \frac{1}{2}\left(H_{1,k-1} + BQ_{11,k-1} - RQ_{11,k-1}\left|Q_{11,k-1}\right| + H_{3,k-1} - BQ_{22,k-1} + RQ_{22,k-1}\left|Q_{22,k-1}\right|\right)$$
$$-\frac{1}{8}B\lambda_2\sqrt{B^2{\lambda_2}^2 + 8\left(H_{1,k-1} + BQ_{11,k-1} - RQ_{11,k-1}\left|Q_{11,k-1}\right| + H_{3,k-1} - BQ_{22,k-1} + RQ_{22,k-1}\left|Q_{22,k-1}\right|\right)}$$

[4.17]

Similarly flow and pressure at node 3 are determined as:

$$Q_{22,k} = \left(H_{3,k} + H_{2,k-1} + BQ_{21,k-1} - RQ_{21,k-1}\left|Q_{21,k-1}\right|\right)\Big/B\,, \qquad [4.18]$$

$$-Q_{31,k} = \left(-H_{3,k} + H_{4,k-1} - BQ_{32,k-1} + RQ_{32,k-1}\left|Q_{32,k-1}\right|\right)\Big/B\,, \qquad [4.19]$$

$$-Q_{L3} = -\lambda_3\sqrt{H_{3,k}}\,, \qquad [4.20]$$

$$H_{3,k} = \frac{1}{8}B^2{\lambda_3}^2 + \frac{1}{2}\left(H_{2,k-1} + BQ_{21,k-1} - RQ_{21,k-1}\left|Q_{21,k-1}\right| + H_{4,k-1} - BQ_{32,k-1} + RQ_{32,k-1}\left|Q_{32,k-1}\right|\right)$$
$$-\frac{1}{8}B\lambda_3\sqrt{B^2{\lambda_3}^2 + 8\left(H_{2,k-1} + BQ_{21,k-1} - RQ_{21,k-1}\left|Q_{21,k-1}\right| + H_{4,k-1} - BQ_{32,k-1} + RQ_{32,k-1}\left|Q_{32,k-1}\right|\right)}$$

[4.21]

## 4.3 Model Verification and Code Development

### 4.3.1 TransAM Software and Code Development

TransAM is a transient modeling software, developed at the University of Toronto, and used globally for modeling distribution systems. It is used in the development of new systems and the examination of existing systems to locate areas where system pressures may exceed limits and to test operational practices. The analyst enters the number of pipes, their lengths and properties, how the pipes are connected at nodes, the expected steady state conditions at each node and the boundary conditions that exist at the nodes. A wide variety of boundary conditions and advanced options exist for the analyst to choose from. The system is capable of modeling large distribution systems and it is use is recorded and verified within many publications [51] [52] [53] [29] [54] [55].

The code that was developed for the work of this thesis may be found in Appendix C. The developed code is capable of modeling basic systems. Similar to TransAM, it takes inputs of pipe lengths, properties, nodal connections and initial estimates of flow and pressure at the nodes. It performs an initial steady state analysis and then uses the output of this analysis to start the transient simulation. Much work went into the development of a code that may be easily reconfigured for different systems, but the code is still limited by the fact that it does not consider wave speed adjustment and boundary conditions are modeled separately. Since the goal of this research was not the development of new commercial software, the code presented in Appendix C contains some logic used in commercial code but may not be used for the simulation of complex systems.

### 4.3.2 Model Verification – A Comparison to TransAM Software

To examine the value of the proposed equation set, the results of a commercial software simulation were compared to that of the proposed equation set. The systems' response to the closure of the downstream valve was used for the comparison.

A 20 second valve closure was modeled by both TransAM and the developed equation set. The response for both models, taken as the hydraulic head at the valve, is shown in Figure 4.3.

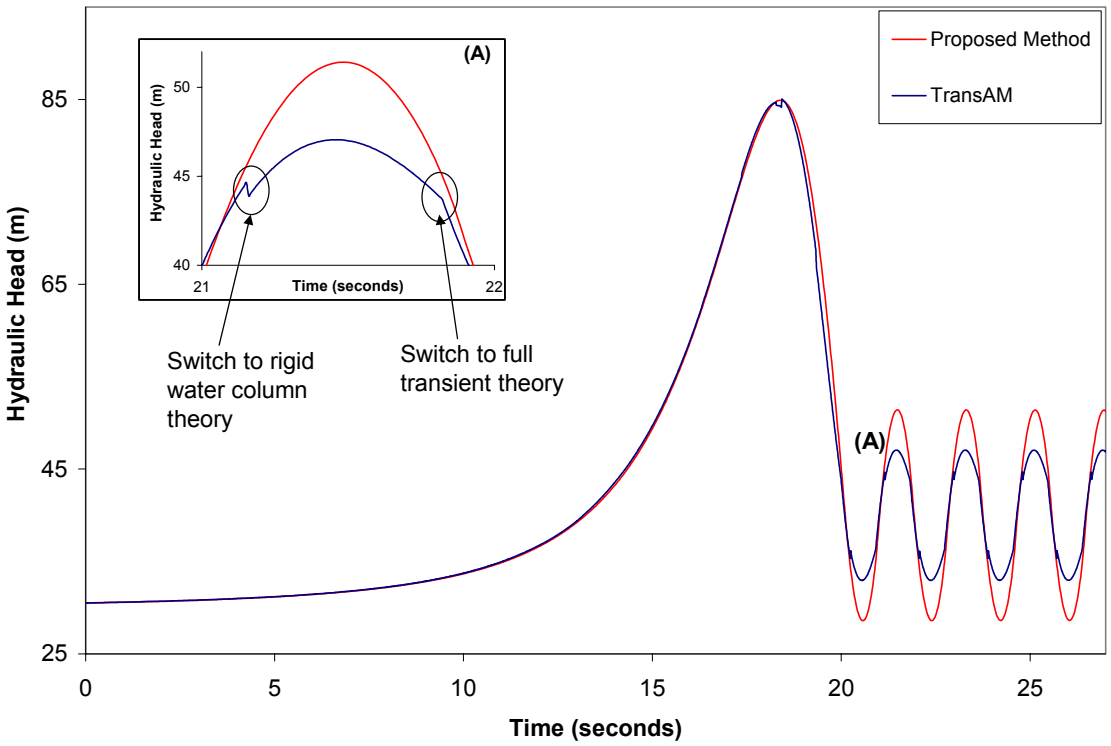


**Figure 4. 3: Comparison of Proposed Method Results with TransAM Results for a 20 Second Valve Closure**

Both models predict the same steady state conditions and initial transient wave front. However, the developed equation set appears to overestimate the dynamics of the hydraulic head following the initial wave front. This can be seen in the exploded view of Figure 4.3. The reason for the discrepancy between the two models is due to a difference in solution methods when fluid acceleration is minimal. The developed equation set is a fully transient solution where compressibility effects are modeled at all times. TransAM uses an adaptive model which switches from a full transient model to a rigid water column model when compressibility effects are not considered significant [56]. This switching effect is accomplished by determining the ratio of the total change in internal energy to the total change in kinetic energy. The ratio, referred to as the compressibility

index, “provides a natural index of the importance of compressibility effects” and by switching to rigid water column theory when compressibility effects are considered negligible, “a forty percent reduction in simulation execution time” may be obtained [56] [57]. The rigid water column model uses a time step that is double that of the normal time step and thus a faster solution may be obtained by switching to this method whenever the effects of compressibility are deemed small. The switching effect is seen by the discontinuities in the pressure trace at the peaks and valleys of the TransAM pressure wave in Figure 4.3. An exploded view better depicts the switching regions. Therefore TransAM’s adaptive model acts to over-damp the transients. This trade off between numerical accuracy and computational speed is justified by the fact that large distribution systems, which may be modeled using TransAM, take a considerable amount of computing effort and therefore computational efficiency is a great importance.

The proposed method offers a good representation of the transient behaviour within the pipeline. The use of an adaptive model within this thesis is not considered necessary since the systems modeled are small and computational time is not a factor. If however, computational time was a factor and an adaptive model was selected, the effect of the switching point discontinuities on the Kalman Filter’s performance would have to be investigated.

## 4.4 The Filter Model and State Space Representation

A state space representation of the describing equations is need for filter implementation. The formulation described above becomes tedious when transformed into a state space representation. Equation set [4.13], [4.14], and [4.17], describing pressure and flow for the inner nodes with leakage, contain the square root valve equation, or equation [4.15]. In the above representation, the unknown leakage states are areas and the leakage flow rate is given by equation [4.15]. Upon transformation to state space, equations [4.13], [4.14], and [4.17] become:

$$H_{2,k} = \frac{1}{8}B^2\lambda_{2,k-1}{}^2 + \frac{1}{2}\left(H_{1,k-1} + BQ_{11,k-1} - RQ_{11,k-1}\left|Q_{11,k-1}\right| + H_{3,k-1} - BQ_{22,k-1} + RQ_{22,k-1}\left|Q_{22,k-1}\right|\right)$$
$$-\frac{1}{8}B\lambda_{2,k-1}\sqrt{B^2\lambda_{2,k-1}{}^2 + 8\left(H_{1,k-1} + BQ_{11,k-1} - RQ_{11,k-1}\left|Q_{11,k-1}\right| + H_{3,k-1} - BQ_{22,k-1} + RQ_{22,k-1}\left|Q_{22,k-1}\right|\right)}$$
[4.22]

$$Q_{12,k} = \frac{1}{8}B\lambda_{2,k-1}{}^2 + \frac{1}{2}\left(\frac{1}{B}H_{3,k-1} - Q_{22k-1} + \frac{R}{B}Q_{22,k-1}\left|Q_{22,k-1}\right|\right) + \frac{3}{2}\left(\frac{1}{B}H_{1,k-1} + Q_{11,k-1} - \frac{R}{B}Q_{11,k-1}\left|Q_{11,k-1}\right|\right)$$
$$-\frac{1}{8}\lambda_{2,k-1}\sqrt{B^2\lambda_{2,k-1}{}^2 + 8\left(H_{1,k-1} + BQ_{11,k-1} - RQ_{11,k-1}\left|Q_{11,k-1}\right| + H_{3,k-1} - BQ_{22,k-1} + RQ_{22,k-1}\left|Q_{22,k-1}\right|\right)}$$
[4.23]

$$Q_{21,k} = \frac{1}{8}B\lambda_{2,k-1}{}^2 + \frac{1}{2}\left(\frac{1}{B}H_{1,k-1} + Q_{11,k-1} - \frac{R}{B}Q_{11,k-1}\left|Q_{11,k-1}\right| - \frac{1}{B}Q_{3,k-1} + Q_{22,k-1} - \frac{R}{B}Q_{22,k-1}\left|Q_{22,k-1}\right|\right)$$
$$-\frac{1}{8}\lambda_{2,k-1}\sqrt{B^2\lambda_{2,k-1}{}^2 + 8\left(H_{1,k-1} + BQ_{11,k-1} - RQ_{11,k-1}\left|Q_{11,k-1}\right| + H_{3,k-1} - BQ_{22,k-1} + RQ_{22,k-1}\left|Q_{22,k-1}\right|\right)}$$
[4.24]

Not only are these equations tedious to program, but filter implementation also requires the partial derivatives of these equations to be taken. Each element of the transition matrix, described in Chapter 3, becomes extremely long and susceptible to programming error. For this reason, the leakage states to be estimated by the filter are flow rates and not areas. This formulation removes equation [4.15] from equation set [4.13], [4.14] and [4.17] and replaces it with a constant flow rate, which is to be estimated. Therefore equation [4.15] becomes:

$$-Q_{L2,k} = -Q_{L2,k-1}, \qquad [4.25]$$

A summation of flow at Node 2 gives:

$$\sum Q = 0 = \frac{-2H_{2,k} + H_{1,k-1} + BQ_{11,k-1} - RQ_{11,k-1}\left|Q_{11,k-1}\right| + H_{3,k-1} - BQ_{22,k-1} + RQ_{22,k-1}\left|Q_{22,k-1}\right|}{B} - Q_{L2,k-1}$$
[4.26]

Finally, solving for $H_{2,k}$ gives:

$$H_{2,k} = \frac{1}{2}\left(H_{1,k-1} + BQ_{11,k-1} - RQ_{11,k-1}\left|Q_{11,k-1}\right| + H_{3,k-1} - BQ_{22,k-1} + RQ_{22,k-1}\left|Q_{22,k-1}\right| - BQ_{L2,k-1}\right). \quad [4.27]$$

Therefore the state space representation of the leakage node equation set is given by equation [4.27] and:

$$Q_{12,k} = \frac{1}{2}\left(H_{1,k-1} + BQ_{11,k-1} - RQ_{11,k-1}\left|Q_{11,k-1}\right| - H_{3,k-1} + BQ_{22,k-1} - RQ_{22,k-1}\left|Q_{22,k-1}\right| + BQ_{L2,k-1}\right)/B\,,$$

[4.28]

$$Q_{21,k} = \frac{1}{2}\left(H_{1,k-1} + BQ_{11,k-1} - RQ_{11,k-1}\left|Q_{11,k-1}\right| - H_{3,k-1} + BQ_{22,k-1} - RQ_{22,k-1}\left|Q_{22,k-1}\right| - BQ_{L2,k-1}\right)/B\,,$$

[4.30]

Equations [4.27]-[4.30] are easily differentiated and programming errors are more easily located due to a reduction in equation length and complexity.

The state vector may be described as:

$$x_j = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \\ x_{10} \\ x_{11} \\ x_{12} \end{bmatrix} = \begin{bmatrix} H_{1,k} \\ H_{2,k} \\ H_{3,k} \\ H_{4,k} \\ Q_{11,k} \\ Q_{12,k} \\ Q_{21,k} \\ Q_{22,k} \\ Q_{31,k} \\ Q_{32,k} \\ Q_{L2,k} \\ Q_{L3,k} \end{bmatrix}. \quad [4.31]$$

The inputs into the system are the reservoir heads and variable valve coefficient $C_v$. A transient response in the system may be generated by a fluctuation in any of the three inputs. Changes in the reservoir heads may represent waves on the reservoir or the fluctuating pressure of a pump if the upstream reservoir is considered as a simplified representation of a constant head pump. The input vector is given as:

$$u_j = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} H_{R1} \\ H_{R2} \\ C_v \end{bmatrix} . \qquad [4.32]$$

In general, the non-linear stochastic difference equation, in state space form, is given as:

$$x_k = f\left(x_{k-1}, u_k, w_{k-1}\right),$$

where $w_{k-1}$, added here, is a set of random variables representing process noise. Present states are written as functions of prior states, present inputs and a random noise variable. To put the above equation set in true state space form, a substitution of one state equation into the other has to be made so that current states are only functions of prior states, plus inputs. The substitution must take place to correctly determine the Jacobian matrix; needed for filter implementation. Rewriting the governing equations in true state space format gives:

$$x_{1,k} = u_{1,k} - \frac{\frac{1}{2}gA^2B^2 + gA^2B\sqrt{B^2 - 4\left(\frac{1+K}{2gA^2}\right)\left(x_{2,k-1} - Bx_{6,k-1} + Rx_{6,k-1}\left|x_{6,k-1}\right| - u_{1,k}\right)}}{(1+k)} - \frac{B^2 - 4\left(\frac{1+K}{2gA^2}\right)\left(x_{2,k-1} - Bx_{6,k-1} + Rx_{6,k-1}\left|x_{6,k-1}\right| - u_{1,k}\right)}{(1+K)} + w_{1,k-1} \qquad [4.33]$$

$$x_{2,k} = \frac{1}{2}\left(x_{1,k-1} + Bx_{5,k-1} - Rx_{5,k-1}\left|x_{5,k-1}\right| + x_{3,k-1} - Bx_{8,k-1} + Rx_{8,k-1}\left|x_{8,k-1}\right| - Bx_{11,k-1}\right) + w_{2,k-1}, \qquad [4.34]$$

$$x_{3,k} = \frac{1}{2}\left(x_{2,k-1} + Bx_{7,k-1} - Rx_{7,k-1}\left|x_{7,k-1}\right| + x_{4,k-1} - Bx_{10,k-1} + Rx_{10,k-1}\left|x_{10,k-1}\right| - Bx_{12,k-1}\right) + w_{3,k-1}, \quad [4.35]$$

$$x_{4,k} = x_{3,k-1} + Bx_{9,k-1} - Rx_{9,k-1}\left|x_{9,k-1}\right| - B\frac{u_{3,k}B + \sqrt{\left(u_{3,k}B\right)^2 + 4u_{3,k}\left(x_{3,k-1} + x_{9,k-1}\left(B - R\left|x_{9,k-1}\right|\right) - u_{2,k}\right)}}{2} + w_{4,k-1}, \quad [4.36]$$

$$x_{5,k} = \frac{-B - \sqrt{B^2 - 4\left(\frac{1+K}{2gA^2}\right)\left(x_{2,k-1} - Bx_{6,k-1} + Rx_{6,k-1}\left|x_{6,k-1}\right| - u_{1,k}\right)}}{1 + K \diagup gA^2} + w_{5,k-1}, \quad [4.37]$$

$$x_{6,k} = \frac{1}{2}\left(x_{1,k-1} + Bx_{5,k-1} - Rx_{5,k-1}\left|x_{5,k-1}\right| - x_{3,k-1} + Bx_{8,k-1} - Rx_{8,k-1}\left|x_{8,k-1}\right| + Bx_{11,k-1}\right)/B + w_{6,k-1}, \quad [4.38]$$

$$x_{7,k} = \frac{1}{2}\left(x_{1,k-1} + Bx_{5,k-1} - Rx_{5,k-1}\left|x_{5,k-1}\right| - x_{3,k-1} + Bx_{8,k-1} - Rx_{8,k-1}\left|x_{8,k-1}\right| - Bx_{11,k-1}\right)/B + w_{7,k-1}, \quad [4.39]$$

$$x_{8,k} = \frac{1}{2}\left(x_{2,k-1} + Bx_{7,k-1} - Rx_{7,k-1}\left|x_{7,k-1}\right| - x_{4,k-1} + Bx_{10,k-1} - Rx_{10,k-1}\left|x_{10,k-1}\right| + Bx_{12,k-1}\right)/B + w_{8,k-1}, \quad [4.40]$$

$$x_{9,k} = \frac{1}{2}\left(x_{2,k-1} + Bx_{7,k-1} - Rx_{7,k-1}\left|x_{7,k-1}\right| - x_{4,k-1} + Bx_{10,k-1} - Rx_{10,k-1}\left|x_{10,k-1}\right| - Bx_{12,k-1}\right)/B + w_{9,k-1}, \quad [4.41]$$

$$x_{10,k} = \frac{u_{3,k}B + \sqrt{\left(u_{3,k}B\right)^2 + 4u_{3,k}\left(x_{3,k-1} + x_{9,k-1}\left(B - R\left|x_{9,k-1}\right|\right) - u_{2,k}\right)}}{2} + w_{10,k-1}, \quad [4.42]$$

$$x_{11,k} = x_{11,k-1} + w_{11,k-1}, \quad [4.43]$$

$$x_{12,k} = x_{12,k-1} + w_{12,k-1}. \qquad [4.44]$$

The output equation is given below:

$$z_k = \begin{bmatrix} x_{1,k} \\ x_{4,k} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} x_k + v_k, \qquad [4.36]$$

where $v_k$ represents measurement noise. The upstream and downstream pipeline heads are taken as outputs, which will also be measurements taken from the system. The state and output equations, derived above, form the state space representation needed for filter implementation.

# Chapter 5: Implementing the Extended Kalman Filter for Leak Detection

## 5.1 Preliminary Remarks

Filter implementation requires the further definition of state, noise and covariance matrices. The transition matrix, or Jacobian of the state equations, is determined. The addition of random error or noise to the plant and measurements is discussed. The initial settings of the covariance matrix $P_k^-$, and the setting and tuning of the plant and measurement covariance matrices $Q_k$ and $R_k$ will be considered.

The Extended Kalman Filter is implemented for the purpose of estimating two "fictitious" leakage states, given as flow rates, within a system. In a pipeline system (in this case a discretized model rather than the real system), leakage is modeled at some specified location. It is not possible for the filter to estimate both the leakage and its location at the same time. Because the filter has information regarding the inlet and outlet head, the filter will be able to "incorrectly" estimate (indeed, be "fooled" into estimating) leakage at two specified points in the filter, the sum of which will be equal to the actual leakage. Further, the location of the true leakage path can then be estimated through interpolation of the known specified locations and the magnitudes of the fictitious leakages. These ideas are described in Section 5.3 of this chapter.

The results of the simulation are given in Section 5.4. The results are analyzed and discussed. The benefits and limitations of this method towards leak detection are discussed.

## 5.2 Filter Implementation

### 5.2.1 Jacobian Matrix Equations

The Jacobian, or partial derivatives of the state equations, is what drives the filtering process. The Jacobian matrix defines the rate of change within the state vector and is needed for determining the *a priori* covariance matrix given by equation [3.34]. The Jacobian is given by equation [3.33] as:

$$J_x = \left.\frac{\partial f[x_k]}{\partial x}\right|_{\hat{x}_k} = \begin{bmatrix} \frac{\partial f[x_{1k}]}{\partial x_1} & \frac{\partial f[x_{1k}]}{\partial x_2} & \cdots \\ \frac{\partial f[x_{2k}]}{\partial x_1} & \frac{\partial f[x_{2k}]}{\partial x_2} & \cdots \\ \vdots & \vdots & \ddots \\ & & \end{bmatrix}_{\hat{x}_k} . \qquad [3.33]$$

Therefore, $J_{2,1}$ is the partial derivative of state equation $x_2$ with respect to $x_1$. A list of the non-zero elements of $J_x$ is given below. In order to more easily understand these equations, the state space representation, $x_k$ and $u_k$, is dropped and head and flow variables are used. The entire equation set, collectively defined as equation [5.1], is given as:

$$J_{1,2} = \frac{-B + \sqrt{B^2 - 3\left(H_{2,k-1} - BQ_{12,k-1} + RQ_{12,k-1}^2 - H_{R1}\right)/\left(gA^2\right)}}{\sqrt{B^2 - 3\left(H_{2,k-1} - BQ_{12,k-1} + RQ_{12,k-1}^2 - H_{R1}\right)/\left(gA^2\right)}},$$

$$J_{1,6} = \frac{\left(-B + \sqrt{B^2 - 3\left(H_{2,k-1} - BQ_{12,k-1} + RQ_{12,k-1}^2 - H_{R1}\right)/\left(gA^2\right)}\right)\left(-B + 2RQ_{12,k-1}\right)}{\sqrt{B^2 - 3\left(H_{2,k-1} - BQ_{12,k-1} + RQ_{12,k-1}^2 - H_{R1}\right)/\left(gA^2\right)}},$$

$$J_{2,1} = J_{2,3} = J_{3,2} = J_{3,4} = J_{6,11} = -J_{7,11} = J_{8,12} = -J_{9,12} = \frac{1}{2},$$

$$J_{2,5} = J_{3,7} = \frac{1}{2}\left(B - 2RQ_{(i)1,k-1}\right) \; where\, i = 1:2,$$

$$J_{i,8} = J_{i,10} = \frac{1}{2}\left(-B + 2RQ_{(i)2,k-1}\right) \; where i = 2:3,$$

$$J_{2,11} = J_{3,12} = -\frac{1}{2}B,$$

$$J_{4,3} = 1 - \frac{BCv}{\sqrt{(BCv)^2 + 4Cv\left(H_{3,k-1} + BQ_{31,k-1} - RQ_{31,k-1}^2 - H_{R2}\right)}},$$

$$J_{4,9} = B - 2RQ_{31,k-1} - \frac{1}{4}\frac{4Cv\left(B - 2RQ_{31,k-1}\right)}{\sqrt{(BCv)^2 + 4Cv\left(H_{3,k-1} + BQ_{31,k-1} - RQ_{31,k-1}^2 - H_{R2}\right)}},$$

$$J_{5,2} = \frac{-1}{\sqrt{B^2 - 3\left(H_{2,k-1} - BQ_{12,k-1} + RQ_{12,k-1}^2 - H_{R1}\right)/\left(gA^2\right)}},$$

$$J_{5,6} = \frac{-B + 2RQ_{12,k-1}}{\sqrt{B^2 - 3\left(H_{2,k-1} - BQ_{12,k-1} + RQ_{12,k-1}^2 - H_{R1}\right)/\left(gA^2\right)}},$$

$$J_{6,1} = -J_{6,3} = J_{7,1} = -J_{7,3} = J_{8,1} = -J_{8,3} = J_{9,1} = -J_{9,3} = \frac{1}{2B},$$

$$J_{6,5} = J_{7,5} = \frac{1}{2B}\left(B - 2RQ_{11,k-1}\right),$$

$$J_{6,8} = J_{7,8} = \frac{1}{2B}\left(B - 2RQ_{22,k-1}\right),$$

$$J_{8,7} = J_{9,7} = \frac{1}{2B}\left(B - 2RQ_{21,k-1}\right),$$

$$J_{8,10} = J_{9,10} = \frac{1}{2B}\left(B - 2RQ_{32,k-1}\right),$$

$$J_{10,3} = \frac{Cv}{\sqrt{\left(BCv\right)^2 + 4Cv\left(H_{3,k-1} + BQ_{31,k-1} - RQ_{31,k-1}^2 - H_{R2}\right)}},$$

$$J_{10,9} = \frac{1}{4}\frac{4Cv\left(B - 2RQ_{31,k-1}\right)}{\sqrt{\left(BCv\right)^2 + 4Cv\left(H_{3,k-1} + BQ_{31,k-1} - RQ_{31,k-1}^2 - H_{R2}\right)}},$$

$$J_{11,11} = J_{12,12} = 1. \qquad [5.1]$$

The measurements taken from the system are pressure heads at the pipe extremes. Therefore the measurement Jacobian, $J_h$, is given as:

$$J_h = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}'. \qquad [5.2]$$

### 5.2.2 Adding Noise to the Simulation

Zero mean noise was added to both the plant and the measurements. Noise is added to the input signal in order to simulate a small transient event at the upstream reservoir of the system. In reality, a constant head reservoir or constant head pump also contain some natural fluctuations due to reservoir waves or pump and motor dynamics that transmit small transients. Measurement noise is added to simulate the uncertainty or error that is associated with head measurement. The addition of noise is illustrated in Figure 5.1.

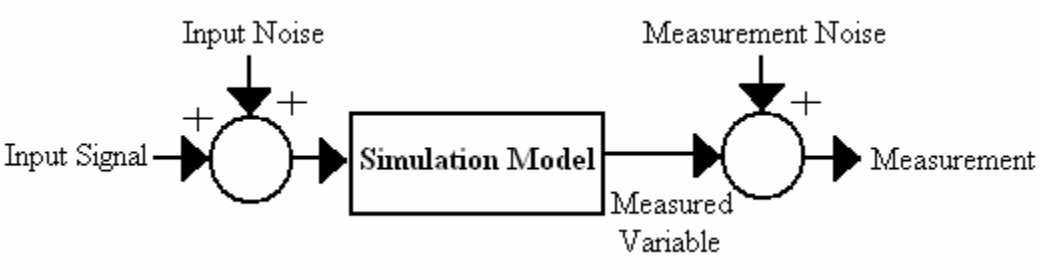


**Figure 5.1: Addition of Noise to the Simulated Plant and Measurement**

### 5.2.3 Initial Conditions and Covariance

The initial conditions needed to start the estimation process include the plant and measurement covariance matrices $Q_k$ and $R_k$, the *a priori* error covariance $P_0^-$, and the state estimate $\hat{x}_0^-$. The initial state estimates were given by a steady state analysis assuming zero leakage. The plant and measurement covariance matrices, $Q_k$ and $R_k$, weight the importance given to the simulated plant and the system measurements. Their setting is not trivial and is discussed below.

The Extended Kalman Filter assumes *a priori* knowledge of both plant and measurement noise and therefore $Q_k$ and $R_k$. Measurement noise is often easily quantified by the precision of the measurement device or sensor. The plant noise is however often difficult to characterize. It may be seen as the deviation between what is happening within the real system and what is simulated or modeled; an uncertainty in the simulated response. In many situations, one or both of these covariance matrices are unknown, however acceptable results may found through trial and error [48].

The measurement covariance, $R_k$, was set by taking into consideration the variance of the random signal added to the simulated measurement. This is akin to setting the measurement noise based on the known resolution of the sensor.

On setting of the plant covariance matrix, $Q_k$, the measurement covariance, $R_k$, was first taken into consideration. The relative magnitudes of both covariance matrices should be similar in order for the filter to recognize the contribution of both measurement and

model. The decision of weight between measurement and plant information would not be as simple in reality as it is in simulation. Equal weighting of both measurement and plant information is, however, a good starting point for all situations. If no noise was assumed in the measurement, i.e. $R_k$ = 0, the state estimates become solely based on measurements. The same holds true for assuming no presence of plant noise. For equal noise contribution, the covariance matrices are given as:

$$R_k = \begin{bmatrix} 20 & 0 \\ 0 & 20 \end{bmatrix}, \quad Q_k = \begin{bmatrix} 20I_4 & 0 & 0 \\ 0 & 1e^{-3}I_6 & 0 \\ 0 & 0 & 1e^{-3}I_2 \end{bmatrix}, \qquad [5.3]$$

where $I_4$ is a 4x4 identity matrix, $I_6$ is a 6x6 identity matrix, and $I_2$ is a 2x2 identity matrix. Note that the three different identity matrices are the size of the head, flow and leakage flow states.

An initial estimate of the *a priori* error covariance is also needed for filter startup. $P_0^-$ is typically set to $P_0^- = CI$, where *I* is an identity matrix and *C* is some large constant [58]. This was the method used in this study. The constant *C* was scaled such that it matched that of the scaling of the plant covariance $Q_k$. This is given as $P_0^- = CIQ_k$. The size of the constant *C* is proportional to the initial rate of convergence of the state estimates. A value of $C = 1e^5$ was used within this thesis. The filter algorithm adjusts the *a priori* error covariance as it runs and therefore the final error covariance will always be similar in magnitude given different initial estimates. The rate of convergence will, however, vary.

## 5.3 Non-Discrete Leak Location and Magnitude Estimates

Discretization, within the method of characteristics, must satisfy the Courant stability condition, given by the Courant number $C_r$. "The Courant number is defined as the ratio of actual wave speed to numerical wave speed", given by the discretization steps *dx/dt*

[15]. For numerical stability to occur, discretization must ensure the Courant number to be $C_r \leq 1$. $C_r = 1$ is the optimal value; smaller values add numerical damping into the simulation. Since the sampling time of measurements is fixed, the spatial discretization step is also fixed and therefore leakage may only be modeled at locations that satisfy the Courant condition. Since it is impractical to expect leakage to occur only at two discrete locations within the pipeline, the estimates of these leakages by the filter would be incorrect and hence are said to be fictitious.

The filtering process produces two discrete fictitious leakage estimates, i.e. two flow rates at two known locations. The sum of the fictitious leakage estimates will be equal to the actual leakage, given only one real leak. If the actual location of the leakage happens at one of these discrete locations, then the magnitude of that leak estimate will be that of the actual leakage, and the other estimate will be zero. Likewise, if there is no leakage within the system, the two fictitious estimates will fluctuate about zero and their sum will be zero. In practical applications, it would be highly unlikely that the location of the leakage will occur at the discrete locations present within the prediction model.

The utility of the fictitious estimates comes from the concept of equivalent systems. Consider Figure 5.2 which displays the head trace for two systems, one with two leaks and one with one leak. Note that from the perspective of the boundaries, these two systems are identical within a steady state analysis.

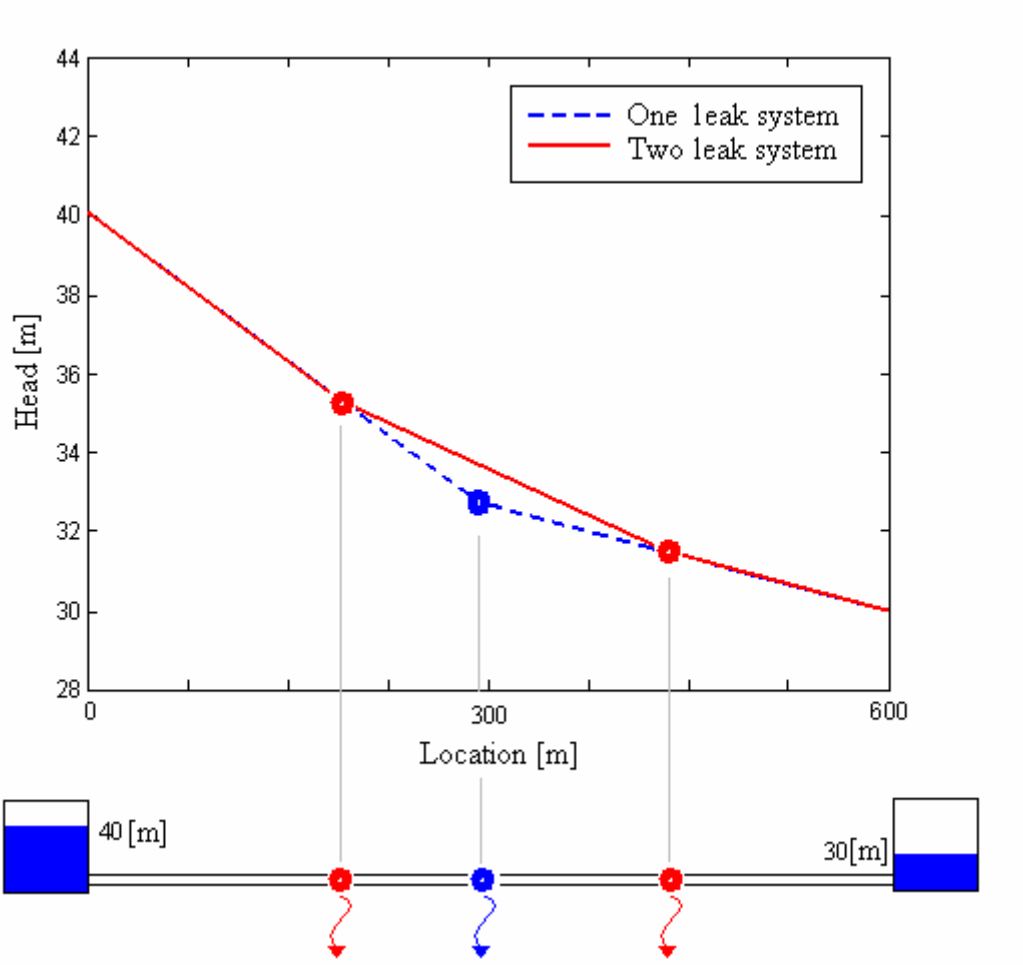


**Figure 5.2: Head Traces for 2 Equivalent Systems**

The following is a derivation for the equations necessary for locating one non-discrete leak, the actual leakage location and magnitude, given the two fixed fictitious filter estimates. Consider Figure 5.3, which depicts two pipes, *pipe a* and *pipe b*, with leakage occurring within them. Assume that *pipe a* is the actual line with leakage, $Q_L$, occurring at any location $x_L$. Now assume that *pipe b* is the prediction model (used by the filter) with leakage, $Q_{L1}$, and $Q_{L2}$, occurring at known locations $x_{L1}$ and $x_{L2}$.

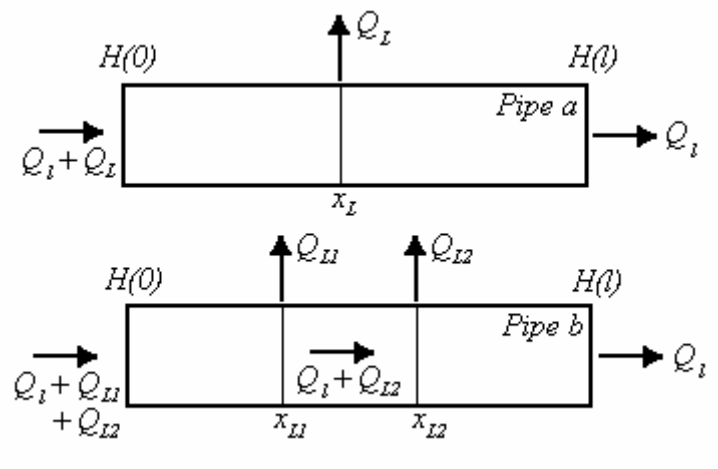


**Figure 5.3: Flow within two Identical Pipelines**

The momentum and continuity equations, which are derived in Chapter 2, represent the flow dynamics occurring within the pipes. These equations are given as:

$$\frac{\partial Q}{\partial t} + gA\frac{\partial H}{\partial x} + \frac{fQ|Q|}{2DA} = 0\,, \qquad [5.4]$$

$$\frac{\partial H}{\partial t} + \frac{a^2}{gA}\frac{\partial Q}{\partial x} = 0\,. \qquad [5.5]$$

For steady state, the temporal terms within equations [5.4] and [5.5] disappear and the equations become:

$$\frac{\partial H}{\partial x} = \frac{fQ|Q|}{2gDA^2}\,, \qquad [5.6]$$

$$\frac{\partial Q}{\partial x} = 0\,. \qquad [5.7]$$

Equation [5.7] states that the mass flow rate at steady state is independent of $x$ and $t$ and is given by the boundary condition at $x = l$. This is given as:

$$Q = Q_l\,. \qquad [5.8]$$

Differentiating equation [5.6] and substituting $Q = Q_l$ gives the Darcy Weisbach equation for steady state flow as:

$$H(x) - H(0) = \frac{fQ_l^2}{2gDA^2} x, \qquad [5.9]$$

where *H(x)* denotes the steady state head at distance *x* from the upstream boundary. The objective is to find $Q_L$ and $x_L$ so that the steady state conditions are the same within both pipes. Continuity states that if the flow is to be the same in both pipes:

$$Q_L = Q_{L1} + Q_{L2}. \qquad [5.10]$$

Applying equation [5.9] across both pipes gives:

$$H(x) - H(0) = \frac{f}{2gDA^2}\left(\left(Q_L + Q_l\right)^2 x_L + Q_l^2\left(l - x_L\right)\right) \; pipe1, \qquad [5.11]$$

$$H(x) - H(0) = \frac{f}{2gDA^2}\left(\left(Q_l + Q_{L1} + Q_{L2}\right)^2 x_{L1} + \left(Q_l + Q_{L2}\right)^2\left(x_{L2} - x_{L1}\right) + Q_l^2\left(l - x_{L2}\right)\right) \; pipe2 \qquad [5.12]$$

For the two pipes to be equivalent, the steady state head drop across the two pipes must be the same. Therefore:

$$\left(Q_L + Q_l\right)^2 x_L + Q_l^2\left(l - x_L\right) = \left(Q_l + Q_{L1} + Q_{L2}\right)^2 x_{L1} + \left(Q_l + Q_{L2}\right)^2\left(x_{L2} - x_{L1}\right) + Q_l^2\left(l - x_{L2}\right). \qquad [5.13]$$

Simplifying equation [5.13] gives:

$$Q_L{}^2 x_L + 2Q_L Q_l x_L = Q_{L1}^2 x_{L1} + Q_{L2}^2 x_{L2} + 2Q_{L1}Q_{L2}x_{L1} + 2Q_{L1}Q_l x_{L1} + 2Q_{L2}Q_l x_{L2}. \qquad [5.14]$$

Dividing both sides by $\frac{1}{Q_l^2}$:

$$\left(\frac{Q_L}{Q_l}\right)^2 x_L + 2\frac{Q_L}{Q_l}x_L = \left(\frac{Q_{L1}}{Q_l}\right)^2 x_{L1} + \left(\frac{Q_{L2}}{Q_l}\right)^2 x_{L2} + 2\frac{Q_{L1}Q_{L2}}{Q_l^2}x_{L1} + 2\frac{Q_{L1}}{Q_l}x_{L1} + 2\frac{Q_{L2}}{Q_l}x_{L2}, \quad [5.15]$$

Assuming that leakage is much smaller than the main flow, $Q_l$, the second order terms in equation [5.15] may be neglected since their magnitude is small compared to the first order terms [32]. Equation [5.15] then becomes:

$$Q_L x_L \approx Q_{L1}x_{L1} + Q_{L2}x_{L2}. \quad [5.16]$$

Therefore, the leak magnitude may be determined by equation [5.10] and its position is given by:

$$x_L \approx \frac{Q_{L1}x_{L1} + Q_{L2}x_{L2}}{Q_L}. \quad [5.17]$$

## 5.4 Simulation Results

The pipeline described in Chapter 4 was simulated using the commercial software MatLAB and the equations developed in Chapter 2. The simulation code may be found within Appendix C. The 600 meter pipeline was discretized into 100 meter sections for the simulation model. (It is to be remembered that the simulation of the pipe line represents the physical pipeline in an applied setting.) Initially the pipeline contains zero leakage to facilitate calculation of steady state conditions. After 90 seconds a leak was added 300 meters from the upstream reservoir. The magnitude of the leakage was taken as approximately 10 percent of the total flow within the pipeline. Since the steady state

flow without leakage was defined to be $Q_{ss}$ = 0.6 [$m^3$/s], the leakage flow was modeled as approximately 0.06 [$m^3$/s].

The upstream and downstream head traces of the pipeline were corrupted by white noise and used as inputs into the Extended Kalman Filter model. Figure 5.4 displays the inputs. The filter model was discretized into 200 meter sections with leak estimates at 200 and 400 meters from the upstream reservoir. Figure 5.5 displays the leakage estimates $Q_{L1}$ and $Q_{L2}$.

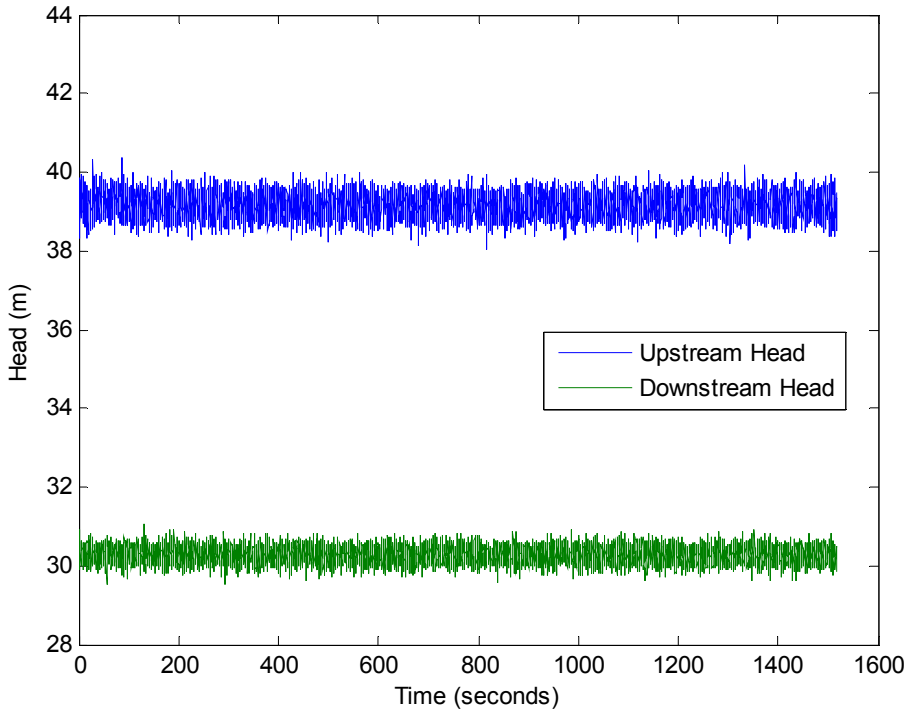


**Figure 5.4: Head Inputs for the Extended Kalman Filter**

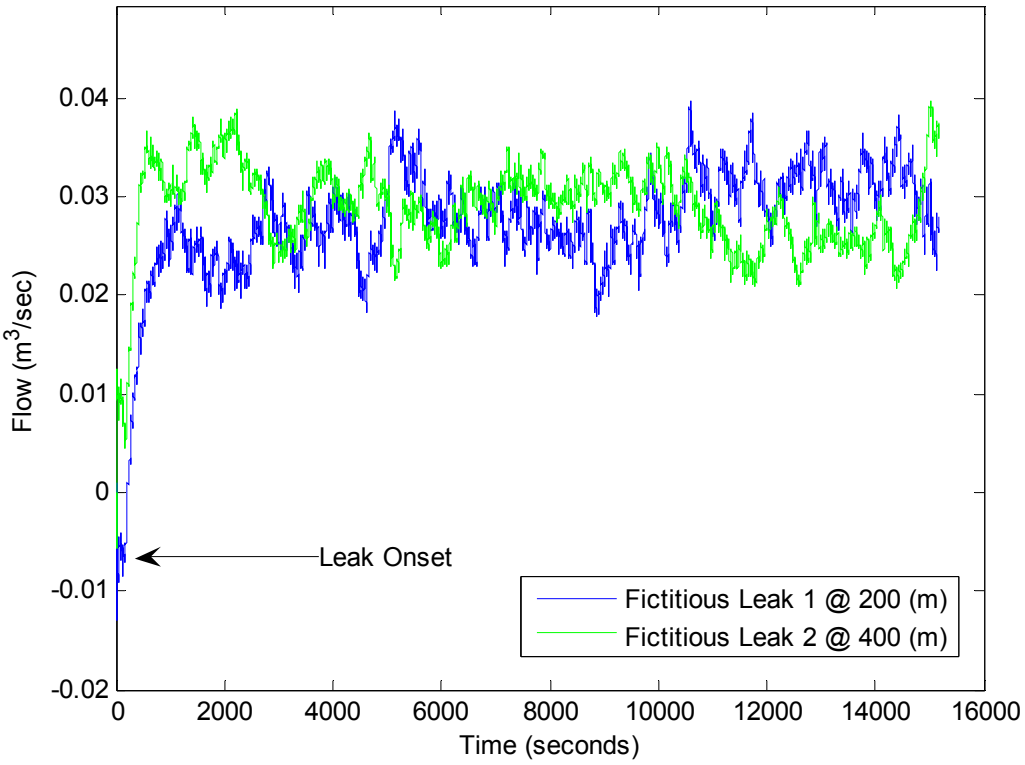


**Figure 5.5: Estimates of the Modeled Leaks, $Q_{L1}$ and $Q_{L2}$**

The modeled fictitious leaks, shown in Figure 5.5, may then be input into equations [5.10] and [5.17] to give the magnitude and position of the actual leakage estimate. Figures 5.6 and 5.7 display the estimated and actual leak magnitude and position respectively.

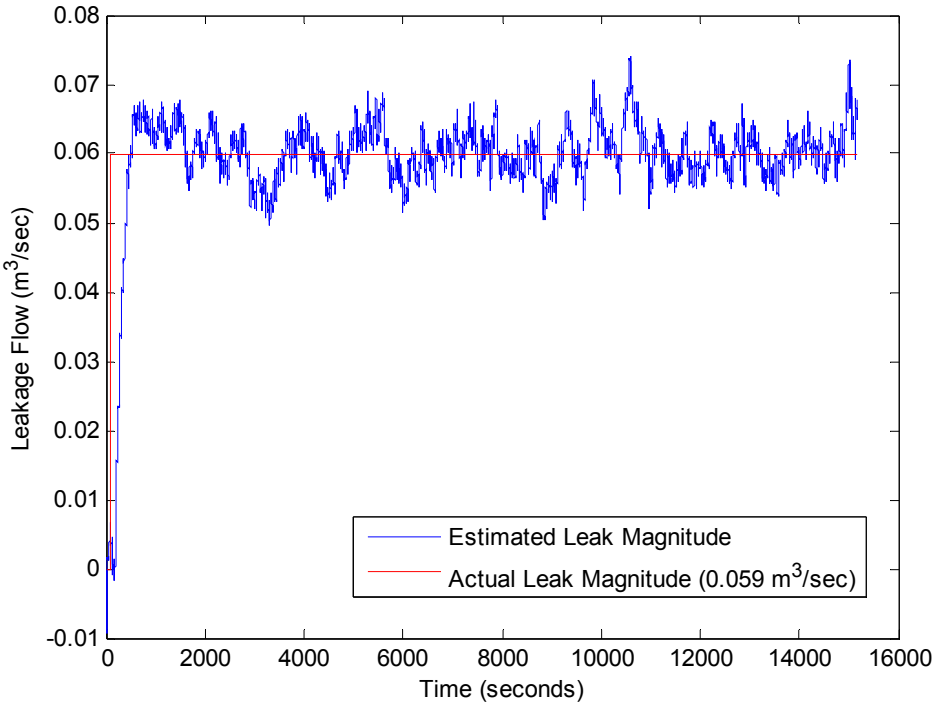


**Figure 5.6: Estimated Leak Magnitude (based on equations [5.10] )**

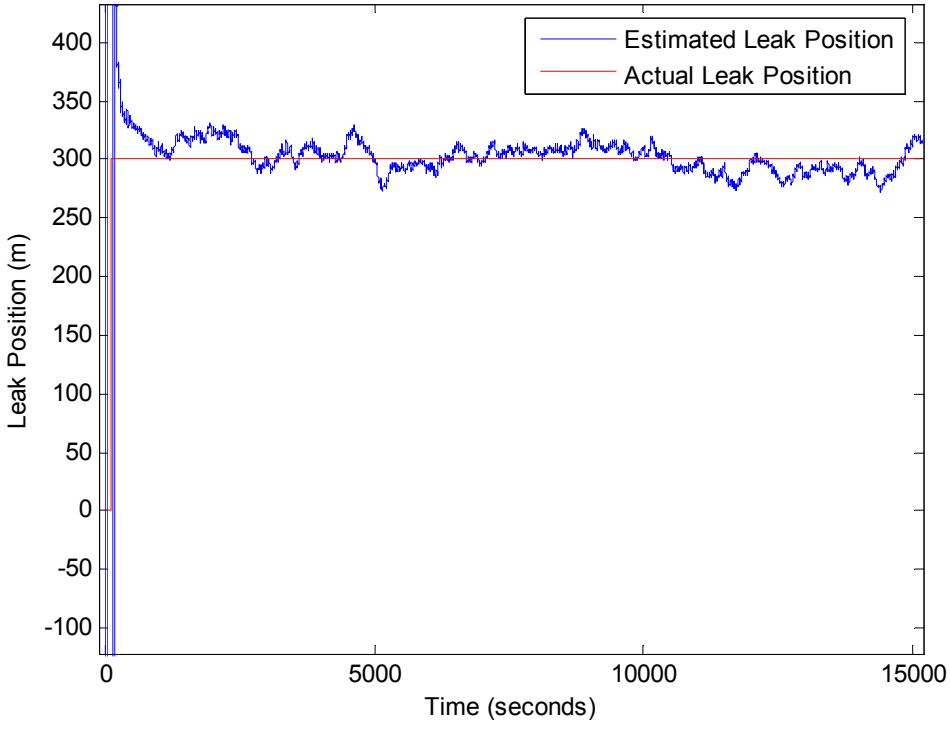


**Figure 5.7: Estimated Leak Position (based on equation [5.17])**

The mean estimates of magnitude and position were calculated to be $Q_L$ = 0.0583 [$m^3/s$] and $x_L$ = 301.4 [m]. These mean values were taken from time = 1200 seconds to the end of the data set. There is a 1.47% error in the mean magnitude estimate and a 0.47% error in the mean position estimate. The standard deviation for the position and magnitude estimates are given as 11.83 [m] and 0.0035 [$m^3/s$] respectively. The filters' ability to locate the position of leakage is further supported by Figure 5.8 below. This figure displays the leak position estimates when leaks were simulated at 100, 200, 300, 400, and 500 meters from the upstream reservoir.

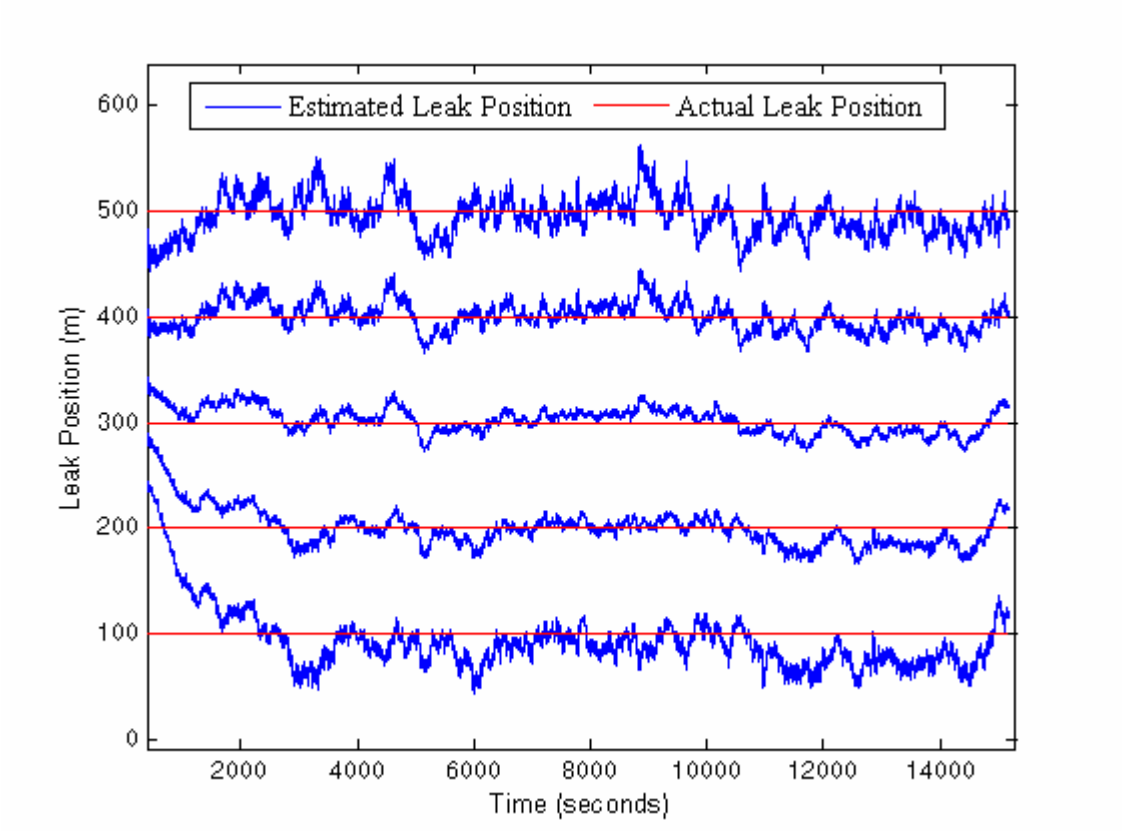


**Figure 5.8: Leak Position Estimates at Varied Positions (based on equation [5.17])**

Table 5.1 displays the mean estimated position and magnitude along with the standard deviation and error in the estimate for each modeled leak position. The average standard deviation, or randomness in the data, was estimated to be 15.09 [m] for the position estimate and 0.0035 [$m^3/s$] for the magnitude estimate.

Note that the standard deviation in the magnitude estimates were all 0.0035 [$m^3/s$], while the standard deviation in the position measurements are similar but not identical. This may be attributed to the fact that equation [5.17], the position estimate equation, is an interpolation between two estimated points. Therefore, when the distance between the actual leakage and the two fixed fictitious estimates increases, which is the case when the leakage lies near the boundaries of the pipeline, the error in the position estimate increases. Equation [5.10], the equation for approximating the leak magnitude, is simply a summation of two direct estimates and therefore will contain only the error given by the Extended Kalman Filtering process.

**Table 5.1: Leak Estimates at Varied Positions**

| Leak Position [m] | Mean Estimated Position [m] | Std [m] | %Error | Actual Magnitude [$m^3/s$] | Mean Estimated Magnitude ($m^3/s$) | Std [$m^3/s$] | %Error |
|---|---|---|---|---|---|---|---|
| 100 | 88.18 | 18.56 | 11.82 | 0.0616 | 0.0602 | 0.0035 | 2.34 |
| 200 | 197.5 | 13.77 | 1.25 | 0.0600 | 0.05879 | 0.0034 | 2.02 |
| 300 | 301.4 | 11.83 | 0.47 | 0.0583 | 0.05745 | 0.0035 | 1.47 |
| 400 | 400.9 | 13.54 | 0.22 | 0.0566 | 0.05615 | 0.0035 | 0.74 |
| 500 | 496.6 | 17.74 | 0.68 | 0.0557 | 0.0549 | 0.0035 | 1.40 |
| | *Average Stats* | 15.09 | 2.89 | | *Average Stats* | 0.0035 | 1.59 |

### 5.4.1 Sensitivity to Leak Magnitude

The accuracy of the leakage estimates depend on the magnitude of the leakage. The minimum magnitude of leakage that may be realized through estimation depends on the resolution of measurement devices. The accuracy of the simulation model will also affect the size of leak that may be detected. In a real system, a model calibration would be done to ensure some level of correlation between measurement and simulation. The degree of fitness between measurement and simulation will give some indication as to the magnitude of the estimates that may be estimated.

Five simulations of leakage at 300 meters from the upstream reservoir were performed. The magnitude of the leakage was varied: 1, 2, 5, 10 and 20% leakage flows were simulated. Table 5.2, figures 5.9 and 5.10 display the results of these estimations.

**Table 5.2: Leak Estimates for Varied Magnitudes**

| % Leakage | Mean Estimated Position [m] | Std [m] | %Error | Actual Magnitude [$m^3/s$] | Mean Estimated Mag.[$m^3/s$] | Std [$m^3/s$] | %Error |
|---|---|---|---|---|---|---|---|
| 1 | 339 | 1.18E+14 | 13.00 | 0.006 | 0.005 | 0.0033 | 16.48 |
| 2 | 317.2 | 347.6 | 5.73 | 0.012 | 0.011 | 0.0032 | 9.52 |
| 5 | 303.9 | 23.9 | 1.30 | 0.030 | 0.030 | 0.0032 | 0.53 |
| 10 | 301.4 | 11.8 | 0.47 | 0.058 | 0.057 | 0.0035 | 1.47 |
| 20 | 306.5 | 6.3 | 2.17 | 0.120 | 0.119 | 0.0031 | 1.08 |

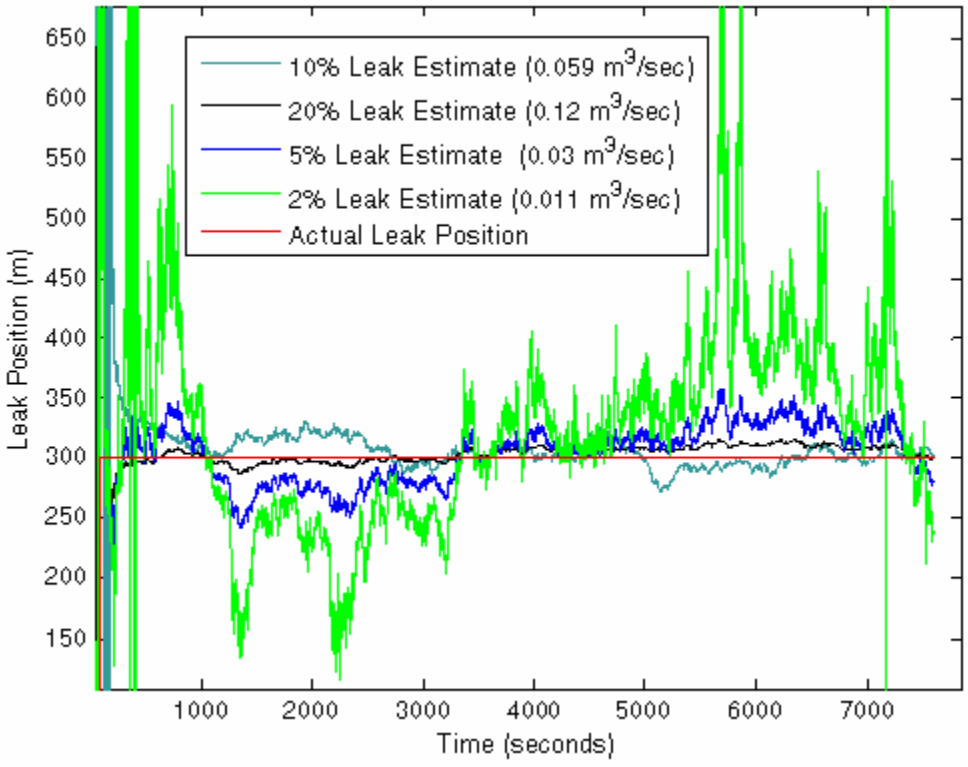


**Figure 5.9: Head Traces for Varied Leak Magnitudes at 300 [m] from Upstream Reservoir**

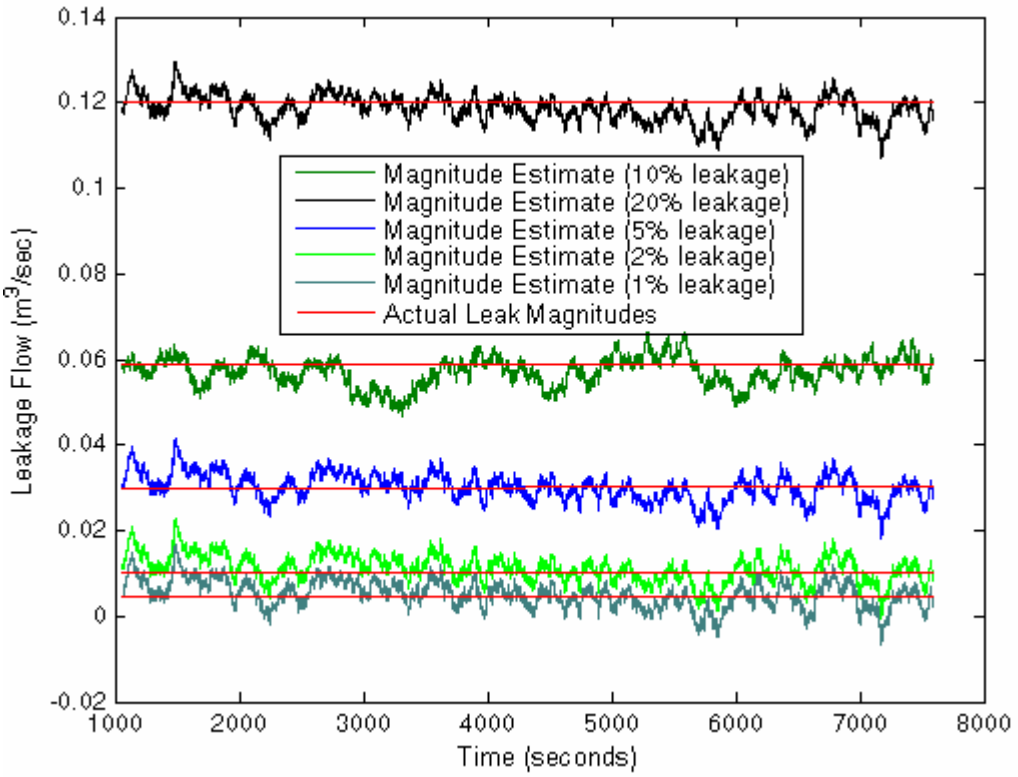


**Figure 5.10: Varying Leak Magnitudes, 1-20% of Steady Flow, at 300 [m] from Upstream Reservoir**

Figure 5.9 and Table 5.2 display that the minimum size leak that can be detected by the filter with some certainty is around 5% of the total flow within the pipeline. The standard deviation of the position estimate for a 5% leak is given as 23.9 [m] whereas the standard deviations for 2% and 1% leaks are given as 347.6 [m] and $1.18e^{14}$ [m] respectively. Note that continuity, equation [5.10], gives an accurate estimate of the leak magnitude even at 1% leakage and the standard deviation remains constant. The first order linear interpolation of the momentum equation, equation [5.17], was not able to determine a reliable estimate of the leakage position smaller than 5% of the total flow, based on the high standard deviation in the estimates.

## 5.5 Proposed Implementation Methods

The following section describes how the leak detection methods of this thesis may be practically applied on real systems. There are a number of areas within water distribution that may apply the methods described in this thesis. The methods are best suited for long

transmission lines or sections of distribution systems that do not contain parallel sections or loops. The methods require pressure measurements, or flow measurements, at the pipe boundaries and therefore the method may be implemented on any section of non-looped pipeline between two access points. Also, the methods are restricted to predicting single leakage zones and therefore these methods may not be applicable to pipelines that contain numerous leaky connections.

Data acquisition should be done at night when consumptive flows are minimal and therefore more predictable. Also because leakage flows will remain the same for a given operating pressure, their relative flow percentage will be at its maximum during the night hours. Therefore it will be easier to detect leakage at night because the consumptive flow variations will be minimized, resulting in minimal modeling error, and leakage flows are more observable at this time.

The methods require an accurate system model. The Extended Kalman filter or other methods may be used for model calibration on a known "representative" section of the line. Consumption rates are also needed. Average nightly consumption rates may be used. It may be beneficial to ask users to restrict usage during the duration of the test. It is also recommended that the data be checked and outliers removed and replace with average values.

The pressure data that is collected needs to be time stamped or synchronized and acquisition must be capable of sampling rates in the order of 0.1 [s]. This of course is relative to the pipe parameters that govern wave speed and the relative complexity, or amount of branches that come off of the transmission line. More users, or branches, will result in the need for a smaller spatial discretization and therefore a quicker sampling rate. Some branches may however be neglected if their night consumption is assumed zero, thus simplify the model and possibly increasing the minimum spatial discretization step needed. Increasing the length of the sampling interval will result in better statistics and a more accurate leak position estimate. Tests should be run over the duration of a couple of hours. A sample size of approximately 3.5 hours was used in this work.

# Chapter 6: Concluding Comments

## 6.1 Project Summary

The objective of this research was to develop and test a model based filtering technique aimed at detecting and therefore reducing leakage flows within water transmission systems. Water distribution systems are notorious for leakage volumes representing on average 30 – 40% and in some cases in excess of 50% of the total volume of pumped water. Effective leak detection must implement a number of the available tools, including steady-state mass volume analysis, transient-based analysis, acoustic analysis, and ground penetrating radar to name a few. Analysts must consider a number of methods because all have their benefits and limitations.

The transient model-based filtering techniques developed in this work are limited to studying single pipe-reaches within which all but the leakage flows are roughly known and pressure measurements are available at the boundaries. This method may be instrumented to encompass an entire system but it is believed that this would be impractical. Instead, this method may be used in conjunction with more general methods, such as mass volume methods, that points out the general regions of leakage within large distribution systems. The analyst may then use these transient methods within "troubled zones" to more accurately determine the positions of leakage.

The Extended Kalman Filter was applied, in simulation, to detect the magnitude and position of leakage within a simulated water transmission line. The estimation process utilized the information transmitted within small transients, originating from the supply reservoir. The transmission line was modeled using the Method of Characteristics, which is the best suited finite-difference solution to the two partial differential equations of momentum and continuity for pipe flow. Inputs for the filtering process were generated through simulation. The upstream and downstream pipeline heads were used as inputs. These measurements were corrupted with white noise with a standard deviation of 0.2

[m], in order to mimic the uncertainty of real measurements. Several simulations were run with different leak positions and magnitudes.

The leak detection process involved first estimating two fictitious leaks, given pressure measurements at the upstream and downstream boundaries of the pipeline. Due to the Courant condition, which constricts discretization within the Method of Characteristics, magnitude and position of leakage could not be simultaneously estimated. Instead, the filter is forced to estimate two "fictitious" leakages at specified locations. The second part of the leak detection process involved applying the concept of equivalent systems to determine the actual position and magnitude of leakage within the pipeline. A relationship, based on the fundamentals of momentum and continuity, was derived for determining the actual estimate of the leak position and magnitude from the two artificially modeled leak states. The magnitude is simply given as the sum of the two leakage estimates and the position is given by a first order location equation, a linear interpolation.

Simulation results show that the estimation process produces quick and accurate estimates of the position and magnitude of leaks as small as 5% of the total flow. The results displayed a linear increase in the standard deviation of the position estimate with a decrease in percent leakage flow, from 6.3 (m) deviation for 20% leak flow to 23.9 (m) deviation for 5% leak flow. The standard deviation in the position estimate then grew exponentially for leak flows between 5% and 1%. When the mean position estimate was calculated for leakage flows ranging from 5% to 20% of the total flow, the average error in the position estimate was approximately 1%. The magnitude of the leakage was estimated with an error of approximately 1% for leakage flows ranging from 5% to 20%. The standard deviation in magnitude estimates remained constant for all positions and leakage flow magnitudes. This was attributed to the fact that the filter model directly estimates the magnitude of the leakage while the position of the leakage was determined after filter estimation through a sort of linear interpolation. Therefore the standard deviation in the magnitude estimate was related to the uncertainty added to the simulated measurements, while the standard deviation of the position magnitude was also related to the uncertainty in the position locating equation, equation [5.17].

The success of this method within a real system relies heavily on model calibration. It is believed that the methods presented in this thesis represent a detection tool that may be successfully implemented upon effective model calibration. Much work and investigation is required in developing and implementing successful loss detection methods.

## 6.2 Conclusions

The model based filtering technique described in this thesis is a tool that could be implemented by itself or in conjunction with other leak detection tools. Based on the results of this study, the following conclusions were drawn.

It may be concluded that the methods described in this thesis present a successful method for quickly identifying leakage. The simulations displayed promising results and were capable of detecting the location of small leakages, 2 percent leakage, within 18 meters. However, this method is limited in its ability since only one leak may be detected, and if multiple leakages occur they may be falsely diagnosed as a single leak.

The methods developed are best suited for long transmission lines or sections of distribution systems that do not contain parallel sections or loops. The methods would be best implemented at night due to two reasons: consumptive flows are more predictable and small, and leakage flows represent a higher percentage of the total flow during the night. Implementation would require the use of time stamped pressure measurements with a capability of sampling at approximately 0.1 [s] over a period of a couple hours. Access points are needed at the boundaries of the test section.

## 6.3 Recommended Future Work

The work of this thesis may be seen as an initial evaluation of implementing Extended Kalman Filtering for leak detection within water transmission lines. Simulations showed

that the Extended Kalman Filter may be used for locating leakage. It is recommended that field tests on a real transmission line should be done to better quantify the usefulness of the filtering process.

An investigation into the use of the Extended Kalman Filter for model calibration purposes is also recommended. The leak detection filter relies on an accurate system model, but the filter itself may be initially used to calibrate the system. The methods of this thesis effectively calibrate for leakage, but the Kalman Filter may also calibrate for friction factors and other quantities that are difficult to identify.

A leak detection method capable of locating multiple leakage zones would be more practical. The methods described in this thesis are only capable of locating one leak within a segment of a transmission line. A study into the use of multiple or banked filtering methods should be done that may be able to differentiate between multiple leaks within a segment of line.

# Reference List


1. A. F Colombo and B. Karney, "Energy and Costs of Leaky Pipes: Toward Comprehensive Picture," Journal of Water Resources Planning and Management vol. 128, 441-450 (2002).

2. C. C. Lai, "Unaccounted for water and the economics of leak detection," Water Supply vol. 9, IR1-1-IR1-8 (1991).

3. K. J. Brothers, "Water leakage and sustainable supply-truth or consequences?," Journal of American Water Works Association vol. 93, 150-152 (2001).

4. B. Bullee, P. Eng Civil Engineer of Water Facilities. November Interview. 2002.

5. M. A. I. Chowdhury, M. F. Ahmed, and M. A. Gaffar, "Water system leak detection in secondary towns of Bangladesh," Water Supply vol. 17, 343-349 (1999).

6. S. Postel, Last Oasis: Facing Water Scarcity (W.W. Norton and Company, New York 1992).

7. E. Miller and G. Tyler, Sustaining the Earth: An Integrated Approach (Thompson Learning, Inc., New York 2002).

8. USGS Water Resources of the United States. Estimated Use of Water in the United States in 2000. Accessed October 2004. http://water.usgs.gov/pubs/circ/2004/circ1268/htdocs/table14.html . 2004.

9. M. Barlow and T. Clarke, Blue Gold: The Fight to Stop the Corporate Theft of the World's Water (The New Press, New York 2002).

10. C Bowden, Killing the Hidden Waters (University of Texas Press, Austin 1977).

11. G. Pelletier, A. Mailhot, and J. Villeneuve, "Modeling Water Pipe Breaks - Three Case Studies," Journal of Water Resources Planning and Management Vol. 129, 115-123 (2003).

12. L. Gohier, Improved Asset Management and Life Cycle Costing of our Critical Underground Infrastructure Systems. February 2, 2004 Seminar.

13. B. Rajani, S. McDonald, and G. Felio, Water Mains Break Data On Different Pipe Materials for 1982 and 1993. 1995. National Research Council of Canada.

14. E. B. Wylie and V. L. Streeter, Fluid Transients (Feb Press, Ann Arbor, Michigan 1983).

15. M. H Chaudhry, Applied Hydraulic Transients (Van Nostrand Reinhold Company Inc., New York 1987).

16. G. Z. Watters, Analysis and Control of Unsteady Flow in Pipelines (Butterworth Publishers, Stoneham, MA 1984).

17. P. Boulos, K. Lansey, and B. Karney, Comprehensive Water Distribution Systems Analysis Handbook: For Engineers and Planners (MWH Soft, Inc., California 2004).

18. O. Hunaidi, and P. Giamou, "Ground-Penetrating Radar for Detection of Leaks in Buried Plastic Water Distribution Pipes." 1998. Lawrence, Kansas, USA, Seventh International Conference on Ground-Penetrating Radar. 5-27-1998.

19. H. V. Fuchs and R. Riehle, "Ten Years of Experience with Leak Detection by Acoustic Signal Analysis," Applied Acoustics vol. 33, 1-19 (1991).

20. B. Brunone, "Transient Test-Based Techniques for Leak Detection in Outfall Pipes," Journal of Water Resources Planning and Management vol. 125, 302-306 (1999).

21. X. Wang, M. F. Lambert, A. R. Simpson, J. A. Liggett, and J. P. Vitkovsky, "Leak Detection in Pipelines using the Damping of Fluid Transients," Journal of Hydraulic Engineering vol. 128, 697-711 (2002).

22. W. Mpesha, S. L. Gassman, and M. H Chaudhry, "Leak Detection in Pipes by Frequency Response Method," Journal of Hydraulic Engineering vol. 127, 134-147 (2001).

23. G. Griebenow and M. Mears, "Leak Detection Implementation: Modeling and Tuning Methods," Journal of Energy Resources Technology vol. 111, 66-72 (1989).

24. C. P. Liou, and J. Tian, "Leak Detection - Transient Flow Simulation Approaches," Journal of Energy Resources Technology vol. 117, 243-248 (1995).

25. C. P. Liou, "Mass Imbalance Error of Waterhammer Equations and Leak Detection," Journal of Fluids Engineering vol. 116, 103-109 (1994).

26. J. Mukherjee and S. Narasimhan, "Leak Detection in Networks of Pipelines by the Generalized Likelihood Ratio Method," Ind. Eng. Chem. Res. vol. 35, 1886-1893 (1996).

27. T. Tucciarelli, A. Criminisi, and D. Termini, "Leak Analysis in Pipeline Systems by Means of Optimal Valve Regulation," Journal of Hydraulic Engineering vol. 125, 277-285 (1999).
28. J. P. Vitkovsky, A. R. Simpson, and M. F. Lambert, "Leak Detection and Calibration Using Transients and Genetic Algorithms," Journal of Water Resources Planning and Management vol. 126, 262-265 (2000).

29. K. W. Tang, B. Brunone, B. Karney, and A. Rossetti, Role and Characterization of Leaks under Transient Conditions. Hotchkiss, R. H and Glade, M. [Building Partnerships - 2000 Joint Conference on Water Resource Engineering and Water Resources Planning & Management]. 7-30-2000.

30. P. Carpentier and G. Cohen, "State estimation and leak detection in water distribution networks," Civil Engineering Systems vol. 8, 247-257 (1991).

31. X. Sun, T. Chen, and H. J. Marquez, "Boiler Leak Detection Using a System Identification Technique," Ind. Eng. Chem. Res. vol. 41, 5447-5454 (2002).

32. A. Benkherouf and A. Y. Allidina, "Leak Detection and Location in Gas Pipelines," IEE Proceedings vol. 135, 142-148 (1988).

33. A. Bhagwat, R. Srinivasan, and P. R. Krishnaswamy, "Fault detection during process transitions: a model-based approach," Chemical Engineering Science vol. 58, 309-325 (2003).

34. D. J. Chmielewski and V. Manousiouthakis, "Loss Accounting and Estimation of Leaks and Instrument Biases Using Time-Series Data," Ind. Eng. Chem. Res. vol. 39, 2336-2344 (2000).

35. D. T. Dalle Molle and D. M. Himmelblau, "Fault Detection in a Single-Stage Evaporator via Parameter Estimation Using the Kalman Filter," Ind. Eng. Chem. Res. vol. 26, 2482-2489 (1987).

36. R. Li and J. H. Olson, "Fault Detection and Diagnosis in a Closed-Loop Nonlinear Distillation Process: Application of Extended Kalman Filters," Ind. Eng. Chem. Res. vol. 30, 898-908 (1991).

37. R. S. Pudar and J. A. Liggett, "Leaks in Pipe Networks," Journal of Hydraulic Engineering vol. 118, 1031-1046 (1992).

38. J. A. Liggett and C. Li-Chung, "Inverse Transient Analysis in Pipe Networks," Journal of Hydraulic Engineering vol. 120, 934-955 (1994).

39. A. S. Willsky, "A Survey of Design Methods for Failure Detection in Dynamic Systems," Automatica vol. 12, 601-611 (1976).

40. R. Isermann, "Process Fault Detection Based on Modeling and Estimation Methods - A Survey," Automatica vol. 20, 387-404 (1983).

41. C. Verde, "Multi-leak Detection and Isolation in Fluid Pipelines," Control Engineering Practice vol. 9, 673-682 (2001).

42. T. Digernes, "Real-time Failure Detection and Identification Aplied to Supervision of Oil Transport in Pipelines," Modeling, identification and control vol. 1, 39-49 (1980).

43. H. Cao, Parameter Estimation Using Extended Kalman Filter for the Swash Plate Assembly and the Control Piston in a Load Sensing Pump. 2001. University of Saskatchewan.

44. E. Kreyszig, Advanced Engineering Mathematics (Wiley, New York 2004).

45. R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," Transaction of the ASME - Journal of Basic Engineering vol. 82 (Series D), 35-45 (1960).

46. H. W. Sorenson, Kalman Filtering Theory and Application (IEEE, 1985).

47. G. Welch, and G. Bishop, An Introduction to the Kalman Filter. 1995. Chapel Hill, NC, USA, University of North Carolina at Chapel Hill.

48. R. G. Brown and P. Y. C. Hwang, Introduction to Random Signals and Applied Kalman Filtering (John Wiley & Sons, Inc., Toronto 1997).

49. G. A. Wright, Parameter Estimation of a Hydraulic Proportional Valve Using Extended Kalman Filtering. 2001. University of Saskatchewan.

50. D. McInnis, B. Karney, and D. H. Axworthy, TransAM Reference Manual. 1998. Ajax, Ontario, HydraTek Associates.

51. B. Karney and M. S. Ghidaoui, "Flexible Discretization Algorithm for Fixed-Grid MOC in Pipelines," Journal of Hydraulic Engineering vol. 123, 1004-1011 (1997).

52. B. Karney and D. McInnis, "Efficient Calculation of Transient Flow in Simple Pipe Networks," Journal of Hydraulic Engineering vol. 118, 1014-1030 (1992).

53. D. McInnis and B. Karney, "Transients in Distribution Networks: Field Tests and Demand Models," Journal of Hydraulic Engineering vol. 121, 218-230 (1995).

54. M. S. Ghidaoui and B. Karney, "Equivalent Differential Equations in Fixed-Grid Characteristics Method," Journal of Hydraulic Engineering vol. 120, 1159-1175 (1994).

55. D. H. Axworthy and B. Karney, "Modeling Low Velocity/High Dispersion Flow in Water Distribution Systems," Journal of Water Resources Planning and Management vol. 122, 218-221 (1996).

56. B. Karney, "Energy Relations in Transient Closed-Conduit Flow," Journal of Hydraulic Engineering vol. 116, 1180-1196 (1990).

57. B. Karney, and D. McInnis, Application of Energy Equations in Unsteady Closed Conduit Flow. V-147-V-165. 1990.

58. L. Ljung and T. Soderstrom, Theory and Practice of Rescursive Identification Cambridge, Massachusetts 1983).

# Appendix A: Discrete State Space Modeling

State space modeling is simply a notational convenience that is used throughout the control and estimation community. An example is used within this section to illustrate state space modeling. Consider the classic example of a mass-spring damper, illustrated in Figure A.1. The describing dynamics of this system are given by the following differential equation:

$$m\frac{d^2}{dt^2}y(t)+c\frac{d}{dt}y(t)+ky(t)=u(t), \qquad [A.1]$$

where *y(t)* is the displacement of the mass, *m*, at time *t*, *c* is the damper constant, *k* is the spring constant and *u(t)* is the input forcing function.

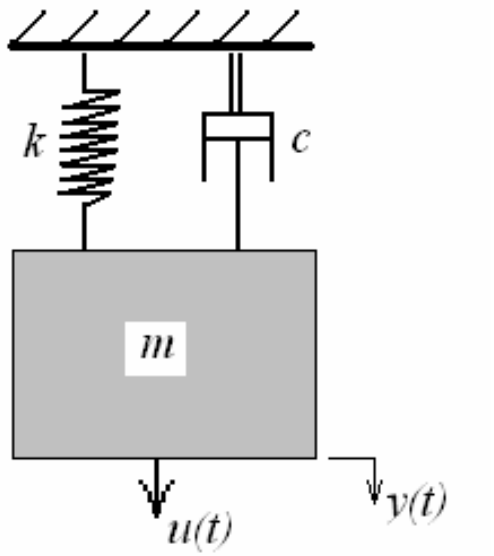


**Figure A.1: Mass Spring Damper System**

If the position as a function of the input forcing function is needed a transfer function approach should be used which is given as:

$$G(s) = \frac{Y(s)}{F(s)} = \frac{1}{ms^2 + cs + k}. \qquad [A.2]$$

Alternatively, if velocity information is also desired a state space approach is most appropriate. In this approach, any variable that may vary temporally or spatially is designated as a state and labelled $x_k$, where the $k$ subscript describes each state from 1 to $k$. If position and velocity are to be tracked, two states $x_1(t)$ and $x_2(t)$ are defined as:

$$x_1 = y(t), \qquad [A.3]$$

and

$$x_2(t) = \frac{d}{dt} y(t) = \frac{dx_1(t)}{dt} = \dot{x}_1(t). \qquad [A.4]$$

Therefore $x_1(t)$ denotes the position and $x_2(t)$ denotes the velocity of the mass. The state space representation is written in such a way that the differential change in one state is given as a function of all prior states. Given equations [A.1], [A.3] and [A.4], the state space representation for this system may be written as:

$$\dot{x}_1(t) = x_2(t),$$

$$\dot{x}_2(t) = -\frac{k}{m} x_1(t) - \frac{c}{m} x_2(t) + \frac{1}{m} u(t),$$

$$y(t) = x_1(t).$$

The state equations are typically written in matrix form. The general state space format is given as:

$$\dot{x}(t) = Ax(t) + Bu(t), \qquad [A.5]$$

$$y(t) = Cx(t) + Du(t). \qquad [A.6]$$

Where, $x(t)$ is the (n x 1) state vector, $A$ is the (n x n) system matrix, $B$ is the (n x r) input matrix, $u(t)$ is the (r x 1) input vector, $y(t)$ is the (p x 1) output vector, $C$ is the (p x n)

output matrix, and $D$ is the (pxr) matrix that represents any coupling between inputs and outputs. Equations [A.5] and [A.6] are known as the state and output equations respectively.

The matrix state space representation is therefore given as:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{c}{m} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} u(t), \qquad \text{[A.7]}$$

$$y(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}. \qquad \text{[A.8]}$$

The state equations are first-order differential equations. The output equations are determined from knowledge of the states and system inputs. In this case and most commonly the $D$ matrix is zero. A different set of state equations may be obtained by performing a transformation on the set of state variables to obtain a new valid set. The number of states present is fixed however by the order of the system. In this case there are two states for the second-order system of a mass-spring damper.

Discrete Time Models

Discrete time models are achieved by discretizing the describing equations. Discrete processes occur in cases where a continuous system is discretely measured. The Kalman Filter and Extended Kalman Filter are most easily applied to discrete state space models and therefore this form was used within the work of this thesis.

Discretization may be accomplished through a number of different methods. The transient equations describing pipeline flow are discretized using a finite-difference method known as the Method of Characteristics. For equations [A.7] and [A.8], describing the dynamics of a mass spring damper, a forward-difference approximation may be used for discretization. The approximation is given as:

$$\dot{x} \approx \frac{x(k+1) - x(k)}{T_S}, \qquad \text{[A.9]}$$

where $T_s$ is the sampling frequency and $k$ refers to a discrete time step. Applying equation [A.9] to [A.7] gives:

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = \left( T_S \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{c}{m} \end{bmatrix} + 1 \right) \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + T_S \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} u(t). \qquad \text{[A.10]}$$

Equation [A.10] may therefore be represented as:

$$x(k+1) = \varphi x(k) + Gu(k), \qquad \text{[A.11]}$$
$$y(k) = Hx(k),$$

where,

$$\varphi = \left( T_S \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{c}{m} \end{bmatrix} + 1 \right),$$

$$G = T_S \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix}, \text{ and}$$

$$H = C.$$

# Appendix B: Probability and Statistics

The following is a basic introduction to probability and random numbers. The Kalman Filter is a tool that provides information from noisy signals. Noise is a random function that can not be described by explicit expressions but may be characterized by probability. Expectation, averages, variance and covariance of random numbers will be covered. These fundamentals are utilized within the development of the Extended Kalman Filter.

Probability

Probability may be seen as the percent chance that one event out of several may occur given a random chance. The probability that a random occurrence will favor a particular event, as in the roll of dice, is given by:

$$p(A) = \frac{Possible\ outcomes\ favoring\ event\ A}{Total\ number\ of\ \ possible\ outcomes}. \qquad [B.1]$$

Probability within Random Variables

Probability within continuous random signals should be thought of in a different way. Random signals are real numbers that may take on any continuous value within the sample space. Since there is no discrete number of outcomes, but rather an infinite amount, the probability of any discrete number $p(A) = 0$. From a limits perspective, this makes sense since one outcome out of an infinite amount is zero. However, probability may be determined for encountering a range of numbers. The cumulative distribution is a function that defines the probability of a random variable. $F_X(x)$ is cumulative because it is the probability of all events up to and including $x$ occurring, or:

$$F_X(x) = p(-\infty, x], \qquad [B.2]$$

where,

$$F_X(x) \to 0 \ \ as\ x \to -\infty,$$

$$F_X(x) \to 1 \ \ as\ x \to \infty.$$

The derivative of the cumulative distribution function is known as the probability density function given as:

$$f_X(x) = \frac{d}{dx} F_X(x), \qquad \text{[B.3]}$$

where

1. $f_X(x)$ is a non-negative function
2. $\int_{-\infty}^{\infty} f_X(x)dx = 1.$

If the probability density function were graphed it would display the distribution of occurrences within the sample space. Therefore the probability over a defined interval is given by integrating the probability density function over that interval. This is given as:

$$p_X[a,b] = \int_a^b f_X(x)dx. \qquad \text{[B.4]}$$

Expectation

The mean or average value of a discrete sample space is given as:

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_N}{N}, \qquad \text{[B.5]}$$

where the sample space has a discrete number of outcomes, $N$. $\overline{X}$ denotes the average of the sample realizations $X_1, X_2, \ldots$. The expected value of $X$ may be described as:

$$Expected\ value\ of\ X = E(X) = \sum_{i=1}^{n} p_i x_i, \qquad \text{[B.6]}$$

where the probability $p_i$ essentially weights the importance of each discrete outcome $x_i$ within the summation of all possible realizations $n$. The continuous definition of expected value is given as:

$$Expected\ value\ of\ X = E(X) = \int_{-\infty}^{\infty} x f_X(x)dx\ . \qquad [B.7]$$

The expectation of a function of $X$ is defined similarly to equations [B.7] as:

$$E\big(g(x)\big) = \int_{-\infty}^{\infty} g(x) f_X(x)dx\ . \qquad [B.8]$$

The expected value is also known as the first statistical moment. Of particular interest is the second moment, in which the variance of the data is derived. It is given as:

$$E\left(X^2\right) = \int_{-\infty}^{\infty} x^2 f_X(x)dx\ . \qquad [B.9]$$

Variance

Define a function that is the random variable minus its expected value $g(x) = X - E(X)$. $g(x)$ is therefore a zero mean function that describes the random behavior of $X$. If the expected value of the second moment of $g(x)$ is taken, the result is the variance of $X$. This is given as:

$$\begin{aligned} Variance\ of\ X = E\left(g(x)^2\right) &= E\left[\left(X - E(X)\right)^2\right] \\ &= E(X^2) - E(X)^2 \end{aligned}\ . \qquad [B.10]$$

Variance describes the amount of dispersion that exists within a random signal. Two signals may have the same expected value but the variance of the two signals may be different. The standard deviation is another way of expressing the dispersion that exists within a random signal. It is given as the square root of the variance, or:

$$Standard\ deviation\ of X = \sigma_X = \sqrt{Variance\ of\ X}\ . \qquad [B.11]$$

Covariance

The covariance of two different variables $X$ and $Y$ is a measure of the correlation or linear dependence that exists between the two variables. The covariance of $X$ and $Y$ is given as:

$$Cov\,of\ X\ and\ Y = \sigma_{XY} = E\left[(X - m_X)(Y - m_Y)\right], \qquad \text{[B.12]}$$

where,

$m_X$ is the mean of $X$, and
$m_Y$ is the mean of $Y$.

Figure B.1 displays two processes that exhibit a high level of linear correlation or covariance and two processes that are linearly uncorrelated.

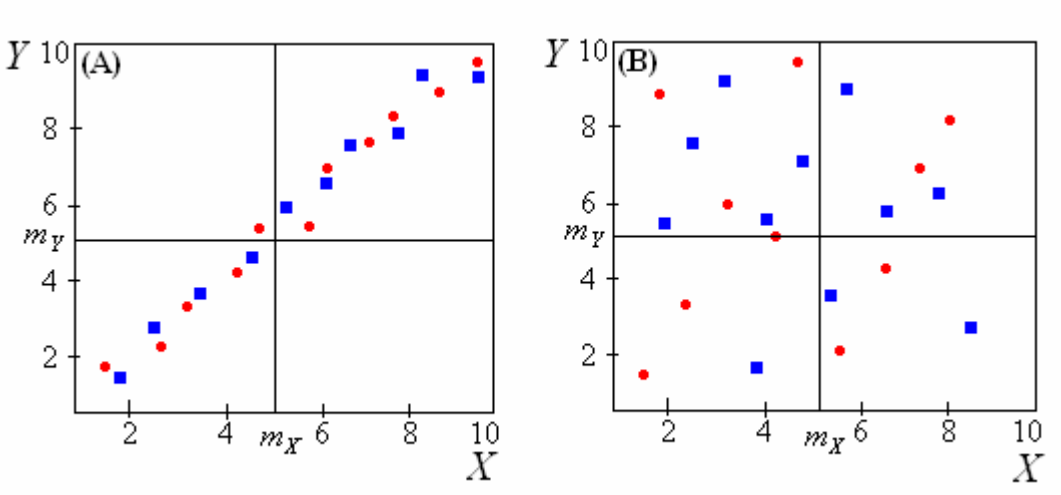


**Figure B.1 : High Covariance (A) Low Covariance (B)**

The two processes $X$ and $Y$ are said to be highly correlated in **(A)** and do not appear to have much correlation within **(B)**. The covariance of **(B)** would be close to zero while the covariance of **(A)** would be close to $\sigma_{XY} = \sigma_X \cdot \sigma_Y$, which is the maximum possible correlation that may exist between processes $X$ and $Y$.

# Appendix C: MatLAB Code

## Simulation Model

```
%Set up Initial Variable Values
ro = 980; % density [kg/m^3]
K = 2.1994e9; % Fluid modulus of elastisity [Pa]
c1 = 1; % constant assuming pipe anchored with expansion joints throughout
g = 9.811; % gravity [m/s^2]

%Set up Node Data
numnode = 7;
node = [1:numnode];
elev = [0,0,0,0,0,0,0]; %node elevation
consmp = [-.62261,0,0,0,0,0,0.62261]; %set the steady state consumption for each node
hgl = [40,38.25,36.5,34.915,33.3,31.665,30]; % Hydraulic grade line
device = [1,2,2,2,2,2,1]; %Boundary condition settings

%Set up Pipedata
numpipe= 6;
minnumreach = 1;
pipe = 1:numpipe;
upstreamnode = [node(1),node(2),node(3),node(4),node(5),node(6)];  %in order [pipe1,pipe2,....pipeN]
downstreamnode = [node(2),node(3),node(4),node(5),node(6),node(7)]; %in order [pipe1,pipe2,....pipeN]
upstreampipe = [pipe(1),pipe(1),pipe(2),pipe(3),pipe(4),pipe(5),pipe(6)]; %set the pipe to the node... in order
[node1,node2,...nodeN]
lengthpipe(1:6) = [100];
f(1:6) = [0.015]; %friction factor in each pipe, assumed constant
[minlength,index] = min(lengthpipe);
diameterpipe(1:6) = [0.5];
E(1:6) = [1.965e11]; %Youngs modulus of elastisity (Steel pipe) [Pa]
e(1:6) = [0.01905];    %Pipe wall thickness [m]
for pipe = 1:numpipe
   a(pipe) = sqrt((K/ro)/(1+((K/E(pipe))*(diameterpipe(pipe)/e(pipe)))*c1)); % Wave speed within the pipe
   A(pipe) = pi/4*diameterpipe(pipe)^2;
   B(pipe) = a(pipe)/(g*A(pipe));     %Wave speed constant used frequently
end

%Set up Time Duration
dt = (minlength/minnumreach)/a(index);
t(1) = 0;
time = 1;

for ryan = 1:300
for pipe = 1:numpipe
   dx(pipe)=a(pipe)*dt;
```

```
   R(pipe) = f(pipe)*dx(pipe)/(2*g*diameterpipe(pipe)*A(pipe)^2); %Resistance constant for each pipe
   j(pipe) = lengthpipe(pipe)/dx(pipe)+1; %Number of grid points in each pipe
   leak(1:numnode) = 0.0;


%Steady State Solver (NEWTONS METHOD)
   LQ(pipe,1:j(pipe)) = (diameterpipe(pipe)^2.5/(0.0826*f(pipe)*lengthpipe(pipe))^0.5)*sqrt(hgl(upstreamnode(pipe))-
(hgl(downstreamnode(pipe))));%-0.5*consmp(1)^2/(2*g*A(pipe)^2)
   LQ1 = LQ;
end %end for pipe = 1:numpipe
   for node = 1:numnode
      uppipes = find(downstreamnode==node); %determine how many pipes are connected to node in question
      downpipes = find(upstreamnode==node);
      if device(node)==1
        consmp(node) = sign(consmp(node))*LQ(upstreampipe(node),1);
      else %if device(d)==2
        if ~isempty(uppipes)
           upQ = 0;
           upQH = 0;
        for Bc1count = 1:length(uppipes)
           upQ = upQ + LQ(uppipes(Bc1count),1);
           upQH = upQH + abs(LQ(uppipes(Bc1count),1)/(hgl(upstreamnode(uppipes(Bc1count)))-
hgl(downstreamnode(uppipes(Bc1count)))));
        end
        else
           upQ = 0;
           upQH = 0;
        end
        if ~isempty(downpipes)
           dwnQ = 0;
           dwnQH = 0;
        for Bc2count = 1:length(downpipes)
           dwnQ = dwnQ + LQ(downpipes(Bc2count),1);
           dwnQH = dwnQH + abs(LQ(downpipes(Bc2count),1)/(hgl(upstreamnode(downpipes(Bc2count)))-
hgl(downstreamnode(downpipes(Bc2count)))));
        end
        else
           dwnQ = 0;
           dwnQH = 0;
        end
        hgl(node) = hgl(node) + ((upQ - dwnQ) - consmp(node))/(.5*(upQH + dwnQH));%this is Newton's METHOD
      end %end if device(node)==1
  end %end for node = 1:numnode
end % for ryan
for pipe = 1:numpipe
   for i = 1
      LH(pipe,i) = (hgl(upstreamnode(pipe)) - (LQ(pipe,1))^2);
```

```
      hgl(upstreamnode(pipe)) = LH(pipe,i);
   end
   for i = 2:j(pipe)-1
      LH(pipe,i) = (hgl(upstreamnode(pipe))-(i-1)*R(pipe)*(LQ(pipe,1))^2);
      LH1 = LH;
   end
   for i = j(pipe)
      LH(pipe,i) = hgl(downstreamnode(pipe)) + (LQ(pipe,j(pipe))^2);
      hgl(downstreamnode(pipe)) = LH(pipe,i);
   end
 end
H = LH;
Q = LQ;
Lleak = leak;


%Valve Stuff
numtimesteps = 100000;
Tau = ones(numnode,numtimesteps);
Ts = 180000; %Start of Valve Closure
leakonset = 1200;
leaky(1:numnode) = 0.0;
leaky(2) = 0.01;
Tc = floor(20/dt); %Time of valve closure (number of time steps)

for time = 1:numtimesteps

if time == leakonset
   Lleak = leaky; % Turn on leak
   leak = leaky;
end

%Calculate Integration Constants
for pipe = 1
   for i = 1
      Bm(pipe,i) = B(pipe);
      Cm(pipe,i) = LH(pipe,i+1) - (B(pipe) - R(pipe)*abs(LQ(pipe,i+1)))*LQ(pipe,i+1);
      Bp(pipe,i) = 0;
      Cp(pipe,i) = 0;
   end
   for i = 2:(j(pipe)-1)
      Bm(pipe,i) = B(pipe);
      Cm(pipe,i) = LH(pipe,i+1) - LQ(pipe,i+1)*(B(pipe) - R(pipe)*abs(LQ(pipe,i+1)));
      Bp(pipe,i) = B(pipe);
      Cp(pipe,i) = LH(pipe,i-1) + LQ(pipe,i-1)*(B(pipe) - R(pipe)*abs(LQ(pipe,i-1)));
      Q(pipe,i) = (Cp(pipe,i) - Cm(pipe,i))/(Bp(pipe,i) + Bm(pipe,i)); %Calculate MOC for Q & H @ interior elements.
      H(pipe,i) = Cp(pipe,i) - Bp(pipe,i)*Q(pipe,i);
```

```
    end
    for i = j(pipe)
        Bp(pipe,i) = B(pipe);
        Cp(pipe,i) = LH(pipe,i-1) + LQ(pipe,i-1)*(B(pipe) - R(pipe)*abs(LQ(pipe,i-1)));
        Bm(pipe,i) = B(pipe);
        Cm(pipe,i) = LH(pipe+1,2) - LQ(pipe+1,2)*(B(pipe) - R(pipe)*abs(LQ(pipe+1,2)));
    end
end
for pipe = 2:numpipe-1
    for i = 1
        Bm(pipe,i) = B(pipe);
        Cm(pipe,i) = LH(pipe,i+1) - (B(pipe) - R(pipe)*abs(LQ(pipe,i+1)))*LQ(pipe,i+1);
        Bp(pipe,i) = B(pipe-1);
        Cp(pipe,i) = LH(pipe-1,j(pipe-1)-1) + LQ(pipe-1,j(pipe-1)-1)*(B(pipe-1) - R(pipe-1)*abs(LQ(pipe-1,j(pipe-1)-1)));
    end
    for i = 2:(j(pipe)-1)
        Bm(pipe,i) = B(pipe);
        Cm(pipe,i) = LH(pipe,i+1) - LQ(pipe,i+1)*(B(pipe) - R(pipe)*abs(LQ(pipe,i+1)));
        Bp(pipe,i) = B(pipe);
        Cp(pipe,i) = LH(pipe,i-1) + LQ(pipe,i-1)*(B(pipe) - R(pipe)*abs(LQ(pipe,i-1)));
        Q(pipe,i) = (Cp(pipe,i) - Cm(pipe,i))/(Bp(pipe,i) + Bm(pipe,i)); %Calculate MOC for Q & H @ interior elements.
        H(pipe,i) = Cp(pipe,i) - Bp(pipe,i)*Q(pipe,i);
    end
    for i = j(pipe)
        Bp(pipe,i) = B(pipe);
        Cp(pipe,i) = LH(pipe,i-1) + LQ(pipe,i-1)*(B(pipe) - R(pipe)*abs(LQ(pipe,i-1)));
        Bm(pipe,i) = B(pipe+1);
        Cm(pipe,i) = LH(pipe+1,2) - LQ(pipe+1,2)*(B(pipe+1) - R(pipe+1)*abs(LQ(pipe+1,2)));
    end
end
for pipe = numpipe
    for i = 1
        Bm(pipe,i) = B(pipe);
        Cm(pipe,i) = LH(pipe,i+1) - (B(pipe) - R(pipe)*abs(LQ(pipe,i+1)))*LQ(pipe,i+1);
        Bp(pipe,i) = B(pipe-1);
        Cp(pipe,i) = LH(pipe-1,j(pipe-1)-1) + LQ(pipe-1,j(pipe-1)-1)*(B(pipe-1) - R(pipe-1)*abs(LQ(pipe-1,j(pipe-1)-1)));
    end
    for i = 2:(j(pipe)-1)
        Bm(pipe,i) = B(pipe);
        Cm(pipe,i) = LH(pipe,i+1) - LQ(pipe,i+1)*(B(pipe) - R(pipe)*abs(LQ(pipe,i+1)));
        Bp(pipe,i) = B(pipe);
        Cp(pipe,i) = LH(pipe,i-1) + LQ(pipe,i-1)*(B(pipe) - R(pipe)*abs(LQ(pipe,i-1)));
        Q(pipe,i) = (Cp(pipe,i) - Cm(pipe,i))/(Bp(pipe,i) + Bm(pipe,i)); %Calculate MOC for Q & H @ interior elements.
        H(pipe,i) = Cp(pipe,i) - Bp(pipe,i)*Q(pipe,i);
    end
    for i = j(pipe)
        Bp(pipe,i) = B(pipe);
```

```
      Cp(pipe,i) = LH(pipe,i-1) + LQ(pipe,i-1)*(B(pipe) - R(pipe)*abs(LQ(pipe,i-1)));
      Bm(pipe,i) = 0;
      Cm(pipe,i) = 0;
   end
end


%Simple and Ordinary One node boundary Conditions
%Calculate Bc and Cc
for node = 1:numnode
   uppipes = find(upstreamnode==node); %determine how many pipes are connected to node in question
   downpipes = find(downstreamnode==node);
   if ~isempty(uppipes)
      upBc(node) = 0;
      for Bc1count = 1:length(uppipes)
         upBc(node) = upBc(node) + 1/Bm(uppipes(Bc1count),1);
      end
      upCc(node) = 0;
      for Cc1count = 1:length(uppipes)
         upCc(node) = upCc(node) + Cm(uppipes(Cc1count),1)/Bm(uppipes(Cc1count),1);
      end %end for loop for Cc1count
   else
      upBc(node) = 0;
      upCc(node) = 0;
   end %end if uppipes
   if ~isempty(downpipes)
      downBc(node) = 0;
      for Bc2count = 1:length(downpipes)
         downBc(node) = downBc(node) + 1/Bp(downpipes(Bc2count),j(downpipes(Bc2count)));
      end
      downCc(node) = 0;
      for Cc2count = 1:length(downpipes)
         downCc(node) = downCc(node) +
Cp(downpipes(Cc2count),j(downpipes(Cc2count)))/Bp(downpipes(Cc2count),j(downpipes(Cc2count)));
      end
   else
      downBc(node) = 0;
      downCc(node) = 0;
   end %end if downpipes
   Bc(node) = (upBc(node) + downBc(node))^-1;
   Cc(node) = Bc(node)*(upCc(node) + downCc(node));

%Boundary Conditions
if device(node) == 1
   for Bc1count = 1:length(uppipes)
      HR = 40 + 0.2*randn(1); %Reservoir Head
if LQ(1,2) > 0
```

```
    Q(uppipes(Bc1count),1) = (-B(pipe) + sqrt(B(pipe)^2 - 4*((1+0.5)/(2*g*A(pipe)^2))*(Cm(uppipes(Bc1count),1) -
HR)))/((1+0.5)/(g*A(pipe)^2));
    H(uppipes(Bc1count),1) = HR - (1 + 0.5)*Q(uppipes(Bc1count),1)^2/(2*g*A(pipe)^2); %Sharp exit minor loss (Energy
equation)
  else
    H(uppipes(Bc1count),1) = HR;
    Q(uppipes(Bc1count),1) = (H(uppipes(Bc1count),1) - Cm(uppipes(Bc1count),1))/Bm(uppipes(Bc1count),1); % from
method of characteristics
  end
  end
  for Bc2count = 1:length(downpipes)
    Cv = (consmp(node)*Tau(node,time))^2/(1*consmp(node)^2/(2*g*A(pipe)^2));%(0.5);
    Q(downpipes(Bc2count),j(downpipes(Bc2count))) = (-Cv*Bc(node)+sqrt((Cv*Bc(node))^2+4*Cv*Cc(node)-
4*Cv*(29.8)))/2;
    H(downpipes(Bc2count),j(downpipes(Bc2count))) = LH(downpipes(Bc2count),j(pipe)-1) +
B(pipe)*LQ(downpipes(Bc2count),j(pipe)-1) - R(pipe)*LQ(downpipes(Bc2count),j(pipe)-
1)*abs(LQ(downpipes(Bc2count),j(pipe)-1)) - B(pipe)*Q(downpipes(Bc2count),j(downpipes(Bc2count)));
  end


 %The Inner Nodes
  else %LEAK NODE
  for Bc2count = 1:length(downpipes)
    H(downpipes(Bc2count),j(downpipes(Bc2count))) = 1/8*B(pipe)^2*Lleak(node)^2 +
1/2*(LH(downpipes(Bc2count),j(downpipes(Bc2count))-1) + B(pipe)*LQ(downpipes(Bc2count),j(downpipes(Bc2count))-1) -
R(pipe)*LQ(downpipes(Bc2count),j(downpipes(Bc2count))-1)*abs(LQ(downpipes(Bc2count),j(downpipes(Bc2count))-1)) +
LH(downpipes(Bc2count)+1,2) - B(pipe)*LQ(downpipes(Bc2count)+1,2) +
R(pipe)*LQ(downpipes(Bc2count)+1,2)*abs(LQ(downpipes(Bc2count)+1,2))) -
1/8*B(pipe)*Lleak(node)*sqrt(B(pipe)^2*Lleak(node)^2 + 8*(LH(downpipes(Bc2count),j(downpipes(Bc2count))-1) +
B(pipe)*LQ(downpipes(Bc2count),j(downpipes(Bc2count))-1) - R(pipe)*LQ(downpipes(Bc2count),j(downpipes(Bc2count))-
1)*abs(LQ(downpipes(Bc2count),j(downpipes(Bc2count))-1)) + LH(downpipes(Bc2count)+1,2) -
B(pipe)*LQ(downpipes(Bc2count)+1,2) + R(pipe)*LQ(downpipes(Bc2count)+1,2)*abs(LQ(downpipes(Bc2count)+1,2))));
  end
  for Bc1count = 1:length(uppipes)
    H(uppipes(Bc1count),1) = 1/8*B(pipe)^2*Lleak(node)^2 + 1/2*(LH(uppipes(Bc1count)-1,j(uppipes(Bc1count)-1)-1) +
B(pipe)*LQ(uppipes(Bc1count)-1,j(uppipes(Bc1count)-1)-1) - R(pipe)*LQ(uppipes(Bc1count)-1,j(uppipes(Bc1count)-1)-
1)*abs(LQ(uppipes(Bc1count)-1,j(uppipes(Bc1count)-1)-1)) + LH(uppipes(Bc1count),2) - B(pipe)*LQ(uppipes(Bc1count),2) +
R(pipe)*LQ(uppipes(Bc1count),2)*abs(LQ(uppipes(Bc1count),2))) - 1/8*B(pipe)*Lleak(node)*sqrt(B(pipe)^2*Lleak(node)^2
+ 8*(LH(uppipes(Bc1count)-1,j(uppipes(Bc1count)-1)-1) + B(pipe)*LQ(downpipes(Bc2count),j(downpipes(Bc2count))-1) -
R(pipe)*LQ(uppipes(Bc1count)-1,j(uppipes(Bc1count)-1)-1)*abs(LQ(uppipes(Bc1count)-1,j(uppipes(Bc1count)-1)-1)) +
LH(uppipes(Bc1count),2) - B(pipe)*LQ(uppipes(Bc1count),2) +
R(pipe)*LQ(uppipes(Bc1count),2)*abs(LQ(uppipes(Bc1count),2))));
    Q(uppipes(Bc1count),1) = (H(uppipes(Bc1count),1) - LH(uppipes(Bc1count),2) + B(pipe)*LQ(uppipes(Bc1count),2) -
R(pipe)*LQ(uppipes(Bc1count),2)*abs(LQ(uppipes(Bc1count),2)))/B(pipe);      Q(uppipes(Bc1count)-1,j(uppipes(Bc1count)-
1)) = Q(uppipes(Bc1count),1) + Lleak(node)*sqrt(abs(H(uppipes(Bc1count),1)))*sign(H(uppipes(Bc1count),1));
end
end
```

```
end

XH1(time,:,:) = H; %Storage of information
XQ1(time,:,:) = Q;
XCv(time) = Cv;
LH = H;
LQ = Q;
Lleak = leak;
t(time+1) = t(time) + dt;
time = time + 1;
end %end of time stepping
pt = t(2:end); %plot timer
```

## Filtering Code

```
%-----------------------------------------------
%                          SETUP
%-----------------------------------------------
clear all;
datasize = 5000;                    %Length of data set to estimate
load('steadystateQandH');
%load('leakat100meterslateonset');
%load('leakat200meterslateonset');
%load('leakat300meterslateonset');
%load('leakat400meterslateonset');
%load('leakat500meterslateonset');
%load('20%leakat300meters');
%load('5%leakat300meters');

ro = 980; % density [kg/m^3]
K = 2.1994e9; % Fluid modulus of elastisity [Pa]
c1 = 1; % constant assuming pipe anchored with expansion joints throughout
g = 9.811; % gravity [m/s^2]

%Set up Node Data
numnode = 7;
node = [1:numnode];
elev = [0,0,0,0,0,0,0]; %node elevation
consmp = [-.62261,0,0,0,0,0,0.62261]; %set the steady state consumption for each node
hgl = [40,38.25,36.5,34.915,33.3,31.665,30]; % Hydraulic grade line
device = [1,2,2,2,2,2,1]; %Boundary condition settings

%Set up Pipedata
numpipe= 6;
minnumreach = 1;
pipe = 1:numpipe;
upstreamnode = [node(1),node(2),node(3),node(4),node(5),node(6)];  %in order [pipe1,pipe2,....pipeN]
```

```
downstreamnode = [node(2),node(3),node(4),node(5),node(6),node(7)]; %in order [pipe1,pipe2,....pipeN]
upstreampipe = [pipe(1),pipe(1),pipe(2),pipe(3),pipe(4),pipe(5),pipe(6)]; %set the pipe to the node... in order
[node1,node2,...nodeN]
lengthpipe(1:6) = [100];
f(1:6) = [0.015]; %friction factor in each pipe, assumed constant
[minlength,index] = min(lengthpipe);
diameterpipe(1:6) = [0.5];
E(1:6) = [1.965e11]; %Youngs modulus of elastisity (Steel pipe) [Pa]
e(1:6) = [0.01905];    %Pipe wall thickness [m]
for pipe = 1:numpipe
   a(pipe) = sqrt((K/ro)/(1+((K/E(pipe))*(diameterpipe(pipe)/e(pipe)))*c1)); % Wave speed within the pipe
   A(pipe) = pi/4*diameterpipe(pipe)^2;
   B(pipe) = a(pipe)/(g*A(pipe));     %Wave speed constant used frequently
end

%Set up Time Duration
dt = (minlength/minnumreach)/a(index);
leak(1:numnode) = 0.0;

%-----------------------------------------------
%                    (Steady state solver)
%-----------------------------------------------
for pipe = 1:numpipe
   dx(pipe)=a(pipe)*dt;
   R(pipe) = f(pipe)*dx(pipe)/(2*g*diameterpipe(pipe)*A(pipe)^2); %Resistance constant for each pipe
   j(pipe) = lengthpipe(pipe)/dx(pipe)+1; %Number of grid points in each pipe
   leak(1:numnode) = 0.0;
end
LH(1:3,1:2) = steadyH;
LQ(1:3,1:2) = steadyQ;
initialH = LH;
initialQ = LQ;

%-----------------------------------------------
%                      Initial Conditions
%-----------------------------------------------

Xest = [LH(1,1);LH(2,1);LH(3,1);LH(3,2);LQ(1,1);LQ(1,2);LQ(2,1);LQ(2,2);LQ(3,1);LQ(3,2);Lleak(2:3)'];
P = eye(12)*1e5;               %Initial error covariance matrix

Qx = zeros(12);
for i = 1:4
   Qx(i,i) = 1;%1e-5;          %noise in head equation
end
for i = 5:10
   Qx(i,i) = 1e-5;%4e-6;          %noise in flow equation
end
```

```
for i = 11:12
   Qx(i,i) = 1e-5;%4e-6;          %noise in leakage estimate
end

P = P.*Qx;        %Adjust initial P matrix according to system noise
                  %matrix (scaling initial error covariance to system
                  %noise matrix)

Rx = eye(2);                %measurement noise covariance matrix
Rx(1,1) = .2;               %noise of upstream head measurement
Rx(2,2) = .2;               %noise of downstream head measurement
%Rx(3,3) = 1;

Hx = zeros(2,12);           %System output matrix
Hx(1,1) = 1;
Hx(2,4) = 1;

F = zeros(12);              %Set up state transistion matrix

%-----------------------------------------------
%                    Parameter Estimation
%-----------------------------------------------
   t(1) = 0;
      iterations = 100000;
   for counter = 1:iterations
% Measurements and Inputs are entered here for each time step
      yact(1,1) = XH1(counter,1,1) + 0.2*randn;
      yact(2,1) = XH1(counter,6,2) + 0.2*randn;
      U1 = 40;
      U2 = 29.8;
      U3 = XCv(counter);
   % The block below defines the linearized state transition matrix (F)

%LH(1,1) Derivatives
for i = 1
F(i,2) = (-B(pipe)+sqrt(B(pipe)^2-3*(LH(2,1)-B(pipe)*LQ(1,2)+R(pipe)*LQ(1,2)^2-U1)/(g*A(pipe)^2)))/(sqrt(B(pipe)^2-
3*(LH(2,1)-B(pipe)*LQ(1,2)+R(pipe)*LQ(1,2)^2-U1)/(g*A(pipe)^2)));
F(i,6) = (-B(pipe)+sqrt(B(pipe)^2-3*(LH(2,1)-B(pipe)*LQ(1,2)+R(pipe)*LQ(1,2)^2-U1)/(g*A(pipe)^2)))*(-
B(pipe)+2*R(pipe)*LQ(1,2))/(sqrt(B(pipe)^2-3*(LH(2,1)-B(pipe)*LQ(1,2)+R(pipe)*LQ(1,2)^2-U1)/(g*A(pipe)^2)));
end
%LH(2,1) Derivatives
for i = 2
F(i,1) = 1/2;
F(i,3) = 1/2;
F(i,5) = 1/2*(B(2) - 2*R(2)*LQ(1,1));
F(i,8) = 1/2*(-B(2) + 2*R(2)*LQ(2,2));
F(i,11) = -1/2*B(2);
```

```
end
%LH(3,1) Derivatives
for i = 3
F(i,2) = 1/2;
F(i,4) = 1/2;
F(i,7) = 1/2*(B(3) - 2*R(3)*LQ(2,1));
F(i,10) = 1/2*(-B(3) + 2*R(3)*LQ(3,2));
F(i,12) = -1/2*B(3);
end
%LH(3,2) Derivatives
for i = 4
F(i,3) = 1-B(pipe)*U3/(sqrt(U3^2*B(pipe)^2+4*U3*LH(3,1)+4*U3*LQ(3,1)*B(pipe)-4*U3*R(pipe)*LQ(3,1)^2-4*U3*U2));
F(i,9) = B(pipe)-2*R(pipe)*LQ(3,1)-1/4*B(pipe)*(4*U3*B(pipe)-
8*U3*R(pipe)*LQ(3,1))/(sqrt(U3^2*B(pipe)^2+4*U3*LH(3,1)+4*U3*LQ(3,1)*B(pipe)-4*U3*R(pipe)*LQ(3,1)^2-4*U3*U2));
end
%LQ(1,1) Derivatives
for i = 5
F(i,2) = -1/(sqrt(B(pipe)^2-3*(LH(2,1)-B(pipe)*LQ(1,2)+R(pipe)*LQ(1,2)^2-U1)/(g*A(pipe)^2)));
F(i,6) = -(-B(pipe)+2*R(pipe)*LQ(1,2))/(sqrt(B(pipe)^2-3*(LH(2,1)-B(pipe)*LQ(1,2)+R(pipe)*LQ(1,2)^2-U1)/(g*A(pipe)^2)));
end
%LQ(1,2) Derivatives
for i = 6
F(i,1) = 1/(2*B(1));
F(i,3) = -1/(2*B(1));
F(i,5) = 1/(2*B(1))*(B(1) - 2*R(1)*LQ(1,1));
F(i,8) = 1/(2*B(1))*(B(2) - 2*R(2)*LQ(2,2));
F(i,11) = 1/2;
end
%LQ(2,1) Derivatives
for i = 7
F(i,1) = 1/(2*B(1));
F(i,3) = -1/(2*B(1));
F(i,5) = 1/(2*B(1))*(B(1) - 2*R(1)*LQ(1,1));
F(i,8) = 1/(2*B(1))*(B(2) - 2*R(2)*LQ(2,2));
F(i,11) = -1/2;
end
%LQ(2,2) Derivatives
for i = 8
F(i,2) = 1/(2*B(2));
F(i,4) = -1/(2*B(2));
F(i,7) = 1/(2*B(2))*(B(2) - 2*R(1)*LQ(2,1));
F(i,10) = 1/(2*B(2))*(B(3) - 2*R(3)*LQ(3,2));
F(i,12) = 1/2;
end
%LQ(3,1) Derivatives
for i = 9
F(i,2) = 1/(2*B(2));
```

```
F(i,4) = -1/(2*B(2));
F(i,7) = 1/(2*B(2))*(B(2) - 2*R(1)*LQ(2,1));
F(i,10) = 1/(2*B(2))*(B(3) - 2*R(3)*LQ(3,2));
F(i,12) = -1/2;
end
%LQ(3,2) Derivatives
for i = 10
F(i,3) = U3/(sqrt(U3^2*B(pipe)^2+4*U3*LH(3,1)+4*U3*LQ(3,1)*B(pipe)-4*U3*R(pipe)*LQ(3,1)^2-4*U3*U2));
F(i,9) = 1/4*(4*U3*B(pipe)-8*U3*R(pipe)*LQ(3,1))/(sqrt(U3^2*B(pipe)^2+4*U3*LH(3,1)+4*U3*LQ(3,1)*B(pipe)-
4*U3*R(pipe)*LQ(3,1)^2-4*U3*U2));
end
%Lleak(2) Derivatives
for i = 11
F(i,11) = 1;
end
%Lleak(3) Derivatives
for i = 12
F(i,12) = 1;
end
%-----------------------------------------------
%   Update P and Calculate Kalman Gain K
%-----------------------------------------------

     P = F*P*F' + Qx;        %Calculate unrefined P matrix
     S = Hx*P*Hx' + Rx;
     K = P*Hx'/S;            %Calculate the Kalman Gain
     %K = ones(12,2)*0.1;
     P = (eye(12) - K*Hx)*P;         %Calculate the refined P matrix

%-------------------------------------------------
% calculate states from original (nonlinear) model
%-------------------------------------------------

 for pipe = 1:numpipe
   for i = 1
      Bm(pipe,i) = B(pipe);
      Cm(pipe,i) = LH(pipe,i+1) - (B(pipe) - R(pipe)*abs(LQ(pipe,i+1)))*LQ(pipe,i+1);
   end
   for i = 2:(j(pipe)-1)
      Bm(pipe,i) = B(pipe);
      Cm(pipe,i) = LH(pipe,i+1) - LQ(pipe,i+1)*(B(pipe) - R(pipe)*abs(LQ(pipe,i+1)));
      Bp(pipe,i) = B(pipe);
      Cp(pipe,i) = LH(pipe,i-1) + LQ(pipe,i-1)*(B(pipe) - R(pipe)*abs(LQ(pipe,i-1)));
      Q(pipe,i) = (Cp(pipe,i) - Cm(pipe,i))/(Bp(pipe,i) + Bm(pipe,i)); %Calculate MOC for Q & H @ interior elements.
      H(pipe,i) = Cp(pipe,i) - Bp(pipe,i)*Q(pipe,i);
   end
   for i = j(pipe)
```

```
      Bp(pipe,i) = B(pipe);
      Cp(pipe,i) = LH(pipe,i-1) + LQ(pipe,i-1)*(B(pipe) - R(pipe)*abs(LQ(pipe,i-1)));
      Bm(pipe,i) = 0;
      Cm(pipe,i) = 0;
   end
end

%Simple and Ordinary One node boundary Conditions
%Calculate Bc and Cc
for node = 1:numnode
   uppipes = find(upstreamnode==node); %determine how many pipes are connected to node in question
   downpipes = find(downstreamnode==node);
   if ~isempty(uppipes)
      upBc(node) = 0;
      for Bc1count = 1:length(uppipes)
         upBc(node) = upBc(node) + 1/Bm(uppipes(Bc1count),1);
      end
      upCc(node) = 0;
      for Cc1count = 1:length(uppipes)
         upCc(node) = upCc(node) + Cm(uppipes(Cc1count),1)/Bm(uppipes(Cc1count),1);
      end %end for loop for Cc1count
   else
      upBc(node) = 0;
      upCc(node) = 0;
   end %end if uppipes
   if ~isempty(downpipes)
      downBc(node) = 0;
      for Bc2count = 1:length(downpipes)
         downBc(node) = downBc(node) + 1/Bp(downpipes(Bc2count),j(downpipes(Bc2count)));
      end
      downCc(node) = 0;
      for Cc2count = 1:length(downpipes)
         downCc(node) = downCc(node) +
Cp(downpipes(Cc2count),j(downpipes(Cc2count)))/Bp(downpipes(Cc2count),j(downpipes(Cc2count)));
      end
   else
      downBc(node) = 0;
      downCc(node) = 0;
   end %end if downpipes
   Bc(node) = (upBc(node) + downBc(node))^-1;
   Cc(node) = Bc(node)*(upCc(node) + downCc(node));

%Boundary Conditions
if device(node) == 1
   for Bc1count = 1:length(uppipes)
  if LQ(1,2) > 0
```

```
    Q(uppipes(Bc1count),1) = (-B(pipe) + sqrt(B(pipe)^2 - 4*((1+0.5)/(2*g*A(pipe)^2))*(Cm(uppipes(Bc1count),1) -
U1)))/((1+0.5)/(g*A(pipe)^2));
    H(uppipes(Bc1count),1) = (U1 - (1 + 0.5)*Q(uppipes(Bc1count),1)^2/(2*g*A(pipe)^2)); %Sharp exit minor loss (Energy
equation)
  else
    H(uppipes(Bc1count),1) = U1;
    Q(uppipes(Bc1count),1) = (H(uppipes(Bc1count),1) - Cm(uppipes(Bc1count),1))/Bm(uppipes(Bc1count),1); % from
method of characteristics
  end
  end
  for Bc2count = 1:length(downpipes)
    Q(downpipes(Bc2count),j(downpipes(Bc2count))) = (-U3*Bc(node)+sqrt((U3*Bc(node))^2+4*U3*Cc(node)-
4*U3*(U2)))/2; %
    H(downpipes(Bc2count),j(downpipes(Bc2count))) = Cc(node) -
Bc(node)*Q(downpipes(Bc2count),j(downpipes(Bc2count)));
  end

 %The Inner Nodes
  else %LEAK NODE
  for Bc2count = 1:length(downpipes)
    H(downpipes(Bc2count),j(downpipes(Bc2count))) = 1/2*(LH(downpipes(Bc2count),j(downpipes(Bc2count))-1) +
B(pipe)*LQ(downpipes(Bc2count),j(downpipes(Bc2count))-1) - R(pipe)*LQ(downpipes(Bc2count),j(downpipes(Bc2count))-
1)*abs(LQ(downpipes(Bc2count),j(downpipes(Bc2count))-1)) + LH(downpipes(Bc2count)+1,2) -
B(pipe)*LQ(downpipes(Bc2count)+1,2) + R(pipe)*LQ(downpipes(Bc2count)+1,2)*abs(LQ(downpipes(Bc2count)+1,2)) -
B(pipe)*Lleak(node));
    Q(downpipes(Bc2count),j(downpipes(Bc2count))) = 0.5*(-LH(downpipes(Bc2count)+1,2) +
B(pipe)*LQ(downpipes(Bc2count)+1,2) - R(pipe)*LQ(downpipes(Bc2count)+1,2)*abs(LQ(downpipes(Bc2count)+1,2)) +
LH(downpipes(Bc2count),j(downpipes(Bc2count))-1) + B(pipe)*LQ(downpipes(Bc2count),j(downpipes(Bc2count))-1) -
R(pipe)*LQ(downpipes(Bc2count),j(downpipes(Bc2count))-1)*abs(LQ(downpipes(Bc2count),j(downpipes(Bc2count))-1)) +
B(pipe)*Lleak(node))/B(pipe);
  end
  for Bc1count = 1:length(uppipes)
    H(uppipes(Bc1count),1) = 1/2*(LH(uppipes(Bc1count)-1,j(uppipes(Bc1count)-1)-1) + B(pipe)*LQ(uppipes(Bc1count)-
1,j(uppipes(Bc1count)-1)-1) - R(pipe)*LQ(uppipes(Bc1count)-1,j(uppipes(Bc1count)-1)-1)*abs(LQ(uppipes(Bc1count)-
1,j(uppipes(Bc1count)-1)-1)) + LH(uppipes(Bc1count),2) - B(pipe)*LQ(uppipes(Bc1count),2) +
R(pipe)*LQ(uppipes(Bc1count),2)*abs(LQ(uppipes(Bc1count),2)) - B(pipe)*Lleak(node));
    Q(uppipes(Bc1count),1) = 1/2*(LH(uppipes(Bc1count)-1,j(uppipes(Bc1count)-1)-1) + B(pipe)*LQ(uppipes(Bc1count)-
1,j(uppipes(Bc1count)-1)-1) - R(pipe)*LQ(uppipes(Bc1count)-1,j(uppipes(Bc1count)-1)-1)*abs(LQ(uppipes(Bc1count)-
1,j(uppipes(Bc1count)-1)-1)) - LH(uppipes(Bc1count),2) + B(pipe)*LQ(uppipes(Bc1count),2) -
R(pipe)*LQ(uppipes(Bc1count),2)*abs(LQ(uppipes(Bc1count),2)) - B(pipe)*Lleak(node))/B(pipe);
    leak(node) = Lleak(node);
end
end
end
    LH = H;
    LQ = Q;
    Lleak = leak;
```

```
    XXest = F*Xest;
    Xest(1) = H(1,1);
    Xest(2:3) = H(1:2,2);
    Xestprime = H(2:3,1);
    Xest(4) = H(3,2);
    Xest(5:6) = Q(1,1:2)';
    Xest(7:8) = Q(2,1:2)';
    Xest(9:10) = Q(3,1:2)';
    Xest(11:12) = leak(2:3)';
    Xest(1:4) = Xest(1:4) + K(1:4,:)*(yact - Hx*Xest);
    Xest(11:12) = Xest(11:12) + K(11:12,:)*(yact - Hx*Xest);
    LH(1,1:2) = Xest(1:2)';
    LH(2,1:2) = Xest(2:3)';
    LH(3,1:2) = Xest(3:4)';
    LQ(1,1:2) = Xest(5:6)';
    LQ(2,1:2) = Xest(7:8)';
    LQ(3,1:2) = Xest(9:10)';
    Lleak(2:3) = Xest(11:12)';

    XH(counter,1,:) = Xest(1:4);
    XQ(counter,1,:) = Xest(5:10);
    Xleak(counter,1,:) = Xest(11:12);
    Xyact(counter,:,:) = yact;
    leakmag(counter) = Xest(11) + Xest(12);
    leakposition(counter) = (200*Xest(11) + 400*Xest(12))/leakmag(counter);
    actmag(counter) = 0;
    actpos(counter) = 0;
    if counter > 600
       actmag(counter) = 0.01*sqrt(36);
       actpos(counter) = 100;
    end
    t(counter+1) = t(counter) + dt;
    pt = t(2:end);


  end
meanpos = mean(leakposition(4000:20000));
meanmag = mean(leakmag(4000:20000));
```